

[illegible]

This Page Blank (uspto)

QY 254 AGCGGACGGAGAGTGGTCACTCTGAGCCAGAGACAGGTTCTTACAAGTCTTGACGGAG 313
 Db 244 AGCGGACCGGAGAGTGGTCACTCTGAGCCAGAGACAGGTTCTTACAAGTCTTGACGGAG 303
 QY 314 AGCAGAAAGCCAAAGGCTTTGAAGGACAGTACAATTTTGAACCATCCAGATGCTTTGATA 373
 Db 304 AGCAGAAAGCCAAAGGCTTTGAAGGACAGTACAATTTTGAACCATCCAGATGCTTTGATA 363
 QY 374 ATGATTTGATGACAGAGCTCTGAGAAACATCGTGAAGGCAAAAAGGTTGAGAGTGGCGGA 433
 Db 364 ATGATTTGATGACAGAGCTCTGAGAAACATCGTGAAGGCAAAAAGGTTGAGAGTGGCGGA 423
 QY 434 CCTATGATTTTGTGACACATCTCAAGGTTACCAAGAGCAACGAGTGTCTTACCTCTGCGAG 493
 Db 424 CCTATGATTTTGTGACACATCTCAAGGTTACCAAGAGCAACGAGTGTCTTACCTCTGCGAG 483
 QY 494 TGGTCTCTTTGAGAGGCACTTTGAGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACC 553
 Db 484 TGGTCTCTTTGAGAGGCACTTTGAGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACC 543
 QY 554 TGGGCTCTCTGAGACACCGAGCTCCGAGCTGCTGAGAGAGTCTCCGGGAG 613
 Db 544 TGGGCTCTCTGAGACACCGAGCTCCGAGCTGCTGAGAGAGTCTCCGGGAG 603
 QY 614 TGGCCGAGGAGGAGCCTGAGAGAGATTTGAGCGAGTACACACTTCTGTGAAGCCGG 673
 Db 604 TGGCCGAGGAGGAGCCTGAGAGAGATTTGAGCGAGTACACACTTCTGTGAAGCCGG 663
 QY 674 CCTTGAGAGAGTCTGCTGCTGCGAGCAAAAGATTA-TGCCGATGTATCATCCAGAGAGA 732
 Db 664 CCTTGAGAGAGTCTGCTGCTGCGAGCAAAAGATTA-TGCCGATGTATCATCCAGAGAGA 723
 QY 733 GTGCAATATGTTGTCATCACTGATCTGTCAGAGCATTCAGAGACATTCGAATGT 792
 Db 724 GTGCAATATGTTGTCATCACTGATCTGTCAGAGCATTCGAATGT 783
 QY 793 GACATCTGCAATGCGACCGAGAGAGGTCAAATGGGCGAGGCTACAGCGGACCTTTTCT 852
 Db 784 GACATCTGCAATGCGACCGAGAGAGGTCAAATGGGCGAGGCTACAGCGGACCTTTTCT 843
 QY 853 GAGCCAGGAGGACCACTCTGGGATGCTGACCTCTGGCAAAAGGTCACATTTGAGTCCAGC 912
 Db 844 GAGCCAGGAGGACCACTCTGGGATGCTGACCTCTGGCAAAAGGTCACATTTGAGTCCAGC 903
 QY 913 AGCAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGGTCCTCCGCGCCGAGATGTGT 972
 Db 904 AGCAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGGTCCTCCGCGCCGAGATGTGT 963
 QY 973 TCAAGGACTGAGGCTGAGGAGCGCCACCAACCCACCTGCTCTCGGCGACCCGAG 1032
 Db 964 TCAAGGACTGAGGCTGAGGAGCGCCACCAACCCACCTGCTCTCGGCGACCCGAG 1023
 QY 1033 GGGAGTGTACAGGAGAGGCTTCTCACTCAAGAGTGAATCTCAGATGTGTCACTCAG 1092
 Db 1024 GGGAGTGTACAGGAGAGGCTTCTCACTCAAGAGTGAATCTCAGATGTGTCACTCAG 1083
 QY 1093 ACTCAACTTGGGAGCACTGACAGGGGCTCTGAGGTTTCAAGCACTTAAGGCTCTGTTG 1152
 Db 1084 ACTCAACTTGGGAGCACTGACAGGGGCTCTGAGGTTTCAAGCACTTAAGGCTCTGTTG 1143
 QY 1153 CGGTTTAAAGATCCCTTAAAGTCACTGAGAAATGCCAGAGATGTGACAGAGCTTGGGA 1212
 Db 1144 CGGTTTAAAGATCCCTTAAAGTCACTGAGAAATGCCAGAGATGTGACAGAGCTTGGGA 1203
 QY 1213 GGCCTTCTGTAGAGATGTGAGGACATTAATTTGGGAAATTAAGAGACAGCCTAGACACT 1272
 Db 1204 GGCCTTCTGTAGAGATGTGAGGACATTAATTTGGGAAATTAAGAGACAGCCTAGACACT 1263
 QY 1273 GGCCTGAGCCTGATGTTTGTGACAGTGAACCAACAGTGGGAGAGGTTTTTCCAGTGTG 1332
 Db 1264 GGCCTGAGCCTGATGTTTGTGACAGTGAACCAACAGTGGGAGAGGTTTTTCCAGTGTG 1323
 QY 1333 ATCTGTTCTTACACTCAGACATTAACAAAAGTTTTGTGAACAAGTACTTCTCTT 1392

Db 1224 ATCTGTTCTTACACTCAGACATTAACAAAAGTTTTGTGAACAAGTACTTCTCTT 1383
 QY 1393 TTTTAAATGTTTAACTATGCTCCATATGTTTCTGTTTCTGTTTCTTAAACAAGGCTGTTG 1452
 Db 1384 TTTTAAATGTTTAACTATGCTCCATATGTTTCTGTTTCTGTTTCTTAAACAAGGCTGTTG 1443
 QY 1453 TGGCCTCAAAACCTTAAATTTTCAATGACCCAGTGGTTTGCAGTCCAGCGGCTTACACGAT 1512
 Db 1444 TGGCCTCAAAACCTTAAATTTTCAATGACCCAGTGGTTTGCAGTCCAGCGGCTTACACGAT 1503
 QY 1513 ATGGGAGCCACTGAGGAGATGTTTTCCCTGCTGTGTGCTTAAAGGAGAGAGAGCA 1572
 Db 1504 ATGGGAGCCACTGAGGAGATGTTTTCCCTGCTGTGTGCTTAAAGGAGAGAGAGCA 1563
 QY 1573 GGGGAGTCCCTGGAACACCCAGATCAACCCAGGCTTGTGCGGGGCCAG 1624
 Db 1564 GGGGAGTCCCTGGAACACCCAGATCAACCCAGGCTTGTGCGGGGCCAG 1615

RESULT 3

This Page Blank (uspto)

Db 754 AACCTGATCGTGCAGACATCCAGACATTCGATGTCATCTGCATGCAATGCAATGCAACCGA 813
Qy 721 GGAGGGTCCAAATGGGGGAGCTCAAGCCGACCTTTTCTAGGCGAGGGACCAACCTCGG 780
Db 814 GGAGGGTCCAAATGGGGGAGCTCAAGCCGACCTTTTCTAGGCGAGGGACCAACCTCGG 873
Qy 781 ATGCTGACCTCTGCAAAAGGTCACATTTGAGTCCAGCAGACCCCACTGA 834
Db 874 ATGCTGACCTCTGCAAAAGGTCACATTTGAGTCCAGCAGACCCCACTGA 927

RESULT 4
US-09-833-381-2048
Sequence 2048, Application US/09833381
Patent No. US20020132090A1
GENERAL INFORMATION:
APPLICANT: Robison, Keith B.
TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2048
LENGTH: 1648
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) ... (1648)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Good clone

Query Match 85.6%; Score 714; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 ATGGCTTCGGGCGGAGCGAAGACCTGAGAGCCCGCGCGAGGCGGACCGTCGCGAC 60
Db 84 ATGGCTTCGGGCGGAGCGAAGACCTGAGAGCCCGCGCGAGGCGGACCGTCGCGAC 143
Qy 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTGAACCTGTGT 120
Db 144 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTGAACCTGTGT 203
Qy 121 GAGAAGTCAATGAGTGTGCTGGAGCAAGAGAGAGTGAAGAGCGGCGGAGAGTGTGTC 180
Db 204 GAGAAGTCAATGAGTGTGCTGGAGCAAGAGAGAGTGAAGAGCGGCGGAGAGTGTGTC 263
Qy 181 ATCTGAGCCAGGACAGGTTCTAACAAGTCTTGAAGGCGGAGAGAGGCGCAAGGCTTTG 240
Db 264 ATCTGAGCCAGGACAGGTTCTAACAAGTCTTGAAGGCGGAGAGAGGCGCAAGGCTTTG 323
Qy 241 AAAGGACAGTAAATTTTGAACCAATCCAGATGCTTGAATTAATTTGATGCAAGAGCT 300
Db 324 AAAGGACAGTAAATTTTGAACCAATCCAGATGCTTGAATTAATTTGATGCAAGAGCT 383
Qy 301 CTGAAGACATCGTGAAGGCGAAGAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
Db 384 CTGAAGACATCGTGAAGGCGAAGAGCGTGAAGTCCGACCTATGATTTTGAACAC 443
Qy 361 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGAGGTGCTTGTGTTGAGGSCATC 420
Db 444 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGAGGTGCTTGTGTTGAGGSCATC 503
Qy 421 TTGGTGTCTACAGCAGGAGATCGGGGACATGTTCCACTGCGCTCTTCGTCGACAC 480
Db 504 TTGGTGTCTACAGCAGGAGATCGGGGACATGTTCCACTGCGCTCTTCGTCGACAC 563
Qy 481 GACTCGAGGTGAGCTGTCTCGAAGATTTCTCGGAGAGTGTGCGGAGGAGGAGCTG 540

This Page Blank (uspto)

614

```
DB 564 GACTCCGACGTCAAGAGTTCTCCGGGACGTCCGAGGAGGACCTG 623
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGGAAGCCGCTTCGAGAGTTCTGCCTG 600
DB 624 GAGCAGATTCTGACGAGTACACCACTTCTGGAAGCCGCTTCGAGAGTTCTGCCTG 683
QY 601 CCGAAGAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCCAT 659
DB 684 CCGAAGAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGCCAT 743
QY 660 CAACCTGATGTCGACGACATCCAGAGATTTGTAATGTCATCTGCAATATGACACCG 719
DB 744 CAACCTGATGTCGACGACATCCAGAGATTTGTAATGTCATCTGCAATATGACACCG 803
QY 720 AGGAGGTCCTCAATGGCGGAGTACAAAGGACCTTTCTGAGCCAGGAGACCACTGG 779
DB 804 AGGAGGTCCTCAATGGCGGAGTACAAAGGACCTTTCTGAGCCAGGAGACCACTGG 863
QY 780 GATGCTGACCTCTGCGCAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 834
DB 864 GATGCTGACCTCTGCGCAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 918
```

```
RESULT 5
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 2041-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 1999-0-20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

dele is good

```
Query Match 48.3%; Score 403; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 24e-207;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 97 GCCAGCGGGAAGTCGACCGTGTGTGAGAGATCATGAGTTGCTGGAGACGAGAGTG 156
DB 41 GCCAGCGGGAAGTCGACCGTGTGTGAGAGATCATGAGTTGCTGGAGACGAGAGTG 100
QY 157 GAAACGCGGAGCGGAGAGTGTGATCTGAGCCAGAGAGTCTTCAAGAGTCTTGACG 216
DB 101 GAAACGCGGAGCGGAGAGTGTGATCTGAGCCAGAGAGTCTTCAAGAGTCTTGACG 160
QY 217 GCAGAGCAGAGGCGCAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATGCTTT 276
DB 161 GCAGAGCAGAGGCGCAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATGCTTT 220
QY 277 GATATGATTTGATGCAACAGACTCTGAGAACTCTGAGGGCAAAAGTGGAGTG 336
DB 221 GATATGATTTGATGCAACAGACTCTGAGAACTCTGAGGGCAAAAGTGGAGTG 280
QY 337 CCGACCTATGATTTTGGAGACACTCAAGTTACAGAGACCAAGTGTCTACCTGCG 396
DB 281 CCGACCTATGATTTTGGAGACACTCAAGTTACAGAGACCAAGTGTCTACCTGCG 340
```

This Page Blank (uspto)

```

RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ. ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Alignment Scores:
Pred. No.: 3,648-173 Length: 1648
Score: 1437.00 Matches: 277
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 99.10% Indels: 1
DB: 10 Gaps: 0

```

This Page Blank (uspto)

US-09-896-522-2 (1-277) x US-09-833-381-2048 (1-1648)

```

OY      1 MetAlaSerAlaGlyGlyGlyAspCysGlySerProAlaProGluValAlaAspArgProHis 20
Db      84 ATGGCTTGGCGGAGGCGAAGACTGCGAGACCCCGCGCGCGAGCCGACCGTCCGCAC 143
OY      21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db      144 CAGCGGCCCTTCTCTATAGGGGTGAGCGCGGCGACTGCCCGGAGGAGTCCGACCGTGT 203
OY      41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db      204 GAGAAAGATCATGAGATTGCTGGGACAGAAACGAGTGAACAGCGGACAGCGAAGGTGCTC 263
OY      61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db      264 ATCCGAGCCAGCAGACAGGTCTTACAAAGTCCCTGACGCGCAGAGCAAGGCCAAGCCCTTG 323
OY      81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      324 AAGGACAGTACAAATTTTGACCATCCAGATGCCCTTGTGATGATGATTGTGACAGACT 383
OY      101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db      384 CTGAAGAACATGCTGAGAGGCGAAAACGGTGGAGGTCCGACCTATGATTTTGTGACACAC 443
OY      121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db      444 TCMAAGTTACCAAGAGACCAAGGTGTCTACCTGCGGACGTGTCTGTGAGGGGCATC 503
OY      141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
Db      504 TTGGTGTCTACAGCCAGAGAGATCCGGACATGTTCCACCTGCGGCTTCTCTGAGGACCC 563
OY      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      564 GACTCCGACGTCAAGCGCTGTCTCGAAGAGTTCGCGGAGCGTCCGCGGAGGAGGACCTG 623
OY      181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db      624 GAGCAGATTCTGACCGACAGTACCAACCTTCGGAAGCCGCGCTTCGAGGAGTCTGCGCTG 683
OY      201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db      684 CCGACAAAGAACTATTGCCGATGTATCATCCACAGAGAGTGGACAAATATGTTGCCAT 743
OY      220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
Db      744 CAACCTGATCGTGACGACATCCAGGACATTCGATGATGTATCTGCAAAATGGCACCG 803
OY      240 GGlYglYSerArgGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlu 260
Db      804 AGGAGGGTCCAAATGGCGGAGCTACAGAGGACCTTTCTGAGCCAGGGGACCAACCTCG 863
OY      260 YMetLeuThrSerGlyLysArgSerHisIleuGlnIleuSerSerArgProHis 277
Db      864 GATGCTGACCTCTGCGAAACGGTCAATTGGAGTCCAGACAGACACCCAC 915

```

RESULT 5

US-10-098-841-53

; Sequence 53, Application US/10098841

; Publication No. US20020197679A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Xu, Chongjun

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunding

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhao, Qing A.

; APPLICANT: Ren, Feiyan

; APPLICANT: Ren, Rui-hong

This Page Blank (uspto)

```
QY 985 CTTGGGGACCCCAACCCCACTGCTTCTCTGGGCAACCCAGGGAGTTAGC 1044
Db 266 CTTGGGGACCCCAACCCCACTGCTTCTCTGGGCAACCCAGGGAGTTAGC 325
QY 1045 AGCGAGGCTTCTCTCACTGAGAGTGAATCTCAGATGTGTCACTCAGACTCAACTTGTCT 1104
Db 326 AGCGAGGCTTCTCTCACTGAGAGTGAATCTCAGATGTGTCACTCAGACTCAACTTGTCT 385
QY 1105 GGGACACTGACAGGCGTCTCTGAGGTTTCAGGCACTTGGCTGCTTGAAGAT 1164
Db 386 GGGACACTGACAGGCGTCTCTGAGGTTTCAGGCACTTGGCTGCTTGAAGAT 445
QY 1165 CCTCTAGTCACTGAGAAATGCAAGATGTGACAGAACTTGGAGGCTTCTGTAG 1224
Db 446 CCTCTAGTCACTGAGAAATGCAAGATGTGACAGAACTTGGAGGCTTCTGTAG 505
QY 1225 GAATG 1229
Db 506 GAATG 510
```

RESULT 8

```
US-09-918-995-30379
; Sequence: 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: MOVIEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

Good

Query Match 26.7%; Score 432.8; DB 11; Length 472;

Best Local Similarity 98.2%; Pred. No. 3,1e-120; Matches 437; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 177 GAGGCGGACCTTCCAGCGGGAAGTGAACCGTGTGTGAGAGATGATGAGTTGCTGG 236
Db 28 GTGGGTGGAANCCGACGCGGGAAGTGAACCGTGTGTGAGAGATGATGAGTTGCTGG 87
QY 237 ACAGAAAGAGGTGAACAGCGGCAAGGAGTGTCTCTTACGCCAGGACAGTTCTA 296
Db 88 ACAGAAAGAGGTGAACAGCGGCAAGGAGTGTCTCTTACGCCAGGACAGTTCTA 147
QY 297 CAAGGTCCTGAGGACAGAGCAAGAGGCAAGGCTTGAAGAAGCAGTACATTTTGACCA 356
Db 148 CAAGGTCCTGAGGACAGAGCAAGAGGCAAGGCTTGAAGAAGCAGTACATTTTGACCA 207
QY 357 TCCAGATGCTTTGATATGATTTGATGACACAGACTCTGAAGACATCGTGAAGGCAA 416
Db 208 TCCAGATGCTTTGATATGATTTGATGACACAGACTCTGAAGACATCGTGAAGGCAA 267
QY 417 AACGTTGAGGTGCCGACCTATATTTTGTGACACTCAAGGTTACAGAGCACCGGT 476
Db 268 AACGTTGAGGTGCCGACCTATATTTTGTGACACTCAAGGTTACAGAGCACCGGT 327
QY 477 GGTCTACCCCTGCGGACGTGTTCTTTTGAAGGCACTTGTGTCTTACAGCCAGAGAT 536
```

This Page Blank (uspto)

597 AAGATTCTCGGAGACGTGCGCGA 621
448 AAGATTCTCGGAGACGTGCGCGA 472

597 AAGATTCTCGGAGACGTGCGCGA 621
448 AAGATTCTCGGAGACGTGCGCGA 472

RESULT 9

331 GCCCAACGGCGGC -GAGCCCTTCTTATAGCGTCAGCGGGAACGCTTAGCGGCAG 388
202 TCGACCGTGTGAGAAATCATGAGATGCTGAGCAAAACGAGGTGAAACAGCGGCAG 261
389 TCTTCGTGTGTAGATCGTCAGCTCTGCGGCAAGATGAGTGAATACTATCGGCAG 448

This page Blank (uspto)

614

```
Db 564 GACTCCGACGTCAGGCTGTCTGGAAGATTCTCCGAGACGTCCGAGGAGGAGACCTG 623
QY 541 GAGCAGATTCTGACGACGATACACCACTTCGTGAAGCCGCTTCGAGAGGTTCTGCTG 600
Db 624 GAGCAGATTCTGACGACGATACCACTTCGTGAAGCCGCTTCGAGAGGTTCTGCTG 683
QY 601 CCGAAGAAAGATATGCGGATGTATCTCCAGAGAGGTGACCAATATGTTGCTCAT 659
Db 684 CCGAAGAAAGATATGCGGATGTATCTCCAGAGAGGTGACCAATATGTTGCTCAT 743
QY 660 CAACCTGATTCGTGACGACATCCAGGACATTTGTAATGTTGACATCTGCAATGACACG 719
Db 744 CAACCTGATTCGTGACGACATCCAGGACATTTGTAATGTTGACATCTGCAATGACACG 803
QY 720 AGGAGGCTCCATGAGCGGAGCTCAAGCGACCTTTCTGAGCCAGGGGACCACTCTG 779
Db 804 AGGAGGCTCCATGAGCGGAGCTCAAGCGACCTTTCTGAGCCAGGGGACCACTCTG 863
QY 780 GATGCTGACCTCTGCGCAAGGTTCACTTTGGAGTCCAGAGACCACTCTG 834
Db 864 GATGCTGACCTCTGCGCAAGGTTCACTTTGGAGTCCAGAGACCACTCTG 918
```

RESULT 5

```
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

date is good
102e

Query Match 48.3%; Score 403; DB 11; Length 472;

Best Local Similarity 100.0%; Pred. No. 2,4e-207; Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 97 GCCAGCGGGAAGTGCACCGTGTGAGAAGATCATGAGTTGCTGGACAGAAAGAGTG 156
Db 41 GCCAGCGGGAAGTGCACCGTGTGAGAAGATCATGAGTTGCTGGACAGAAAGAGTG 100
QY 157 GAACAGCGGCAAGCGGAGGTGTCATCTCTGAGCCAGGACAGGTTCTTCAAGGTCTTGACG 216
Db 101 GAACAGCGGCAAGCGGAGGTGTCATCTCTGAGCCAGGACAGGTTCTTCAAGGTCTTGACG 160
QY 217 GCAAGAGCAAGAGCCCAAGCCCTTGAAGAGACATGTAATTTGACATCCAGATGCTTT 276
Db 161 GCAAGAGCAAGAGCCCAAGCCCTTGAAGAGACATGTAATTTGACATCCAGATGCTTT 220
QY 277 GATAATGATTGATGACAGGACTCTGAAGAACATCGTGAAGGCAAAAAGGTGAGGTG 336
Db 221 GATAATGATTGATGACAGGACTCTGAAGAACATCGTGAAGGCAAAAAGGTGAGGTG 280
QY 337 CCGACCTATGATTGTTGTGACACATCAAGGTTACAGAGACCAAGGTGTTACCTTCG 396
Db 281 CCGACCTATGATTGTTGTGACACATCAAGGTTACAGAGACCAAGGTGTTACCTTCG 340
```

```
QY 397 GACGTGTTCTGTTTGAAGGCAATCTTGATGTTCTACAGCCAGAGACATCCGGACATGTTTC 456
Db 341 GACGTGTTCTGTTTGAAGGCAATCTTGATGTTCTACAGCCAGAGACATCCGGACATGTTTC 400
QY 457 CACCTGCGCTCTTCTGTTGACACCGACCTCCGACGTGAGGCTGT 499
Db 401 CACCTGCGCTCTTCTGTTGACACCGACCTCCGACGTGAGGCTGT 443
```

This Page Blank (uspto)


```

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 504 TTGGTGTCTTACAGCGAGAGATCCGGGACATGTTCCACTCGGCTCTTCTGACAC 563
QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
DB 564 GACTCCGACGTCAGGCTGTCTCGAAGAGCTTCCGGGACGTCCGGCGAGGAGGACCTG 623
QY 181 GlnGlnIleLeuThrGlnTyrThrPheValIleProAlaPheGlnGluPheCysLeu 200
DB 624 GACGCAATTTCAACGACACACACTCTGTGAAGCCGCTTCGAGAGTTCCTGCG 683
QY 201 ProThrTyrSerTyr-AlaAspValIleIleProArgGlyValAspAspMetValAla 220
DB 684 CCGACAAAGAAATATGCGGATGATGATCCACGAGAGTGGACATATGTTGCCAT 743
QY 220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysTyrTrpHisAr 240
DB 744 CAACCTGATCGGACGACATCCAGACATCTGAAATGTGACATCTGCAATGGCACCG 803
QY 240 GGIyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAspHisProG 260
DB 804 AGGAGGGTTCATGGGCGGAGCTCAAGCGACCTTTTCTGAGCCAGGGGACCACTTGG 863
QY 260 YMetLeuThrSerGlyIleArgSerHisLeuGluSerSerSerArgProHis 277
DB 864 GATGCTGACCTCTGGCAACGGTCACTTGTGAGTCCAGCAGACAGCCCCAC 915

```

RESULT 5

```

US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION.
; APPLICANT: (HySeq, Inc)
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 1999-01-20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

```

Alignment Scores:

```

Pred. No.: 5,166-134 Length: 472
Score: 134.00 Matches: 134
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.38% Indels: 0
DB: 11 Gaps: 0

```

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

```

QY 33 AlaSerGlyLysSerThrValCysGluIleMetGluLeuLeuGlyGlnAsnGluVal 52
DB 41 GCCAGCGGGAATCGACCTGTGTGAGAAATCATGATTCCTGGGACAGACAGATG 100
QY 53 GlnGlnArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrIleValLeuThr 72
DB 101 GAACAGCGGACGCGAAGGTGTGATCTGAGCAGACAGGTTCTACAAAGTCTGACG 160
QY 73 AlaGlnGlnIleAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPhe 92

```

```

DB 161 GCAGAGCAGAAAGGCCAAGGCTTGAAGAAGACAGTACAAATTTGACCATCCAGATGCTTT 220
QY 93 AspAsnAspLeuMetHisArgThrIleLeuLysAsnIleValGluGlyIleThrValGluVal 112
DB 221 GATTAATGATTTGATGACAGGACTCTGAAGAACATCTGTGAGGGCAAAACGGGTGAGGTG 280
QY 113 ProThrTyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAla 132
DB 281 CCGACCTAATCATTTTGTGACACACTCAAGTTCACAGAACACACGCTGCTACCTGCG 340
QY 133 AspValValLeuPheGlnGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPhe 152
DB 341 GACGTGCTCTGTTTGAAGGACATCTTGTTCTTACAGCAGAGAGATCCGGGACATGTT 400
QY 153 HisLeuArgLeuPheValAspThrAspSerAspValArgLeu 166
DB 401 CACCTGGCCTCTTGTGACACCGACTCCGACGTCAAGGCTG 442

```

This Page Blank (uspto)

This Page Blank (uspto)

QY 149 CCGACACGCGCCCTTCTGATAGGGGTGAGCGCGGCATCTGCCAGCGGAAGTCGACCG 208
DB 150 CCGACACGCGCGCCCTTCTGATAGGGGTGAGCGCGGCATCTGCCAGCGGAAGTCGACCG 209
QY 209 TGTGTGAGAGATCAATGAGTTGCTGGGACAGAAACGAGGTGAAACAGCGGCAGCGGAGG 268
DB 210 TGTGTGAGAGATCAATGAGTTGCTGGGACAGAAACGAGGTGAAACAGCGGCAGCGGAGG 269
QY 269 TGTGTGAGAGATCAATGAGTTGCTGGGACAGAAACGAGGTGAAACAGCGGCAGCGGAGG 328
DB 270 TGTGTGAGAGATCAATGAGTTGCTGGGACAGAAACGAGGTGAAACAGCGGCAGCGGAGG 329
QY 329 CCTTGAAGAGACATCAATTTTGAACATCCGATGCTTGTGATGATTTGATGACA 388
DB 330 CCTTGAAGAGACATCAATTTTGAACATCCGATGCTTGTGATGATTTGATGACA 389
QY 389 GGAATCTGAAGACATCTGAGAGGCAAAACGCTGAGAGTGCAGACCTATGATTTGTGA 448
DB 390 GGAATCTGAAGACATCTGAGAGGCAAAACGCTGAGAGTGCAGACCTATGATTTGTGA 449
QY 449 CAGACTCAGAGTTACAGAGACCAAGGTGTCTACCTCGAGAGTGTCTCTTTGAGG 508
DB 450 CAGACTCAGAGTTACAGAGACCAAGGTGTCTACCTCGAGAGTGTCTCTTTGAGG 509
QY 509 GCATCTTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGAGG 568
DB 510 GCATCTTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGAGG 569
QY 569 AGACCGAGTCCGAGCGTCAAGGTGTCTCGAAGAGTCTCCGGGACGTGGCCGAGAGGAGG 628
DB 570 AGACCGAGTCCGAGCGTCAAGGTGTCTCGAAGAGTCTCCGGGACGTGGCCGAGAGGAGG 629
QY 629 ACCTGAGAGATTTGAGCGAGTACACCCCTTGTGAAAGCCGCTTCTGAGAGTTCT 688
DB 630 ACCTGAGAGATTTGAGCGAGTACACCCCTTGTGAAAGCCGCTTCTGAGAGTTCT 689
QY 689 GCCTGCGGACAAAGAGTATGCGGATGATCATCCGAGAGTGAACAATATGATTG 748
DB 690 GCCTGCGGACAAAGAGTATGCGGATGATCATCCGAGAGTGAACAATATGATTG 749
QY 749 CCATCAACCTGATCTGAGACATCAGAGACATCTGAAATGTGACATCTGCAATGCGC 808
DB 750 CCATCAACCTGATCTGAGACATCAGAGACATCTGAAATGTGACATCTGCAATGCGC 809
QY 809 ACCGAGAGGGGTCAATGGGCGGAGCTACAGGGGACCTTTCTGAGCCAGGGAGCACCC 868
DB 810 ACCGAGAGGGGTCAATGGGCGGAGCTACAGGGGACCTTTCTGAGCCAGGGAGCACCC 869
QY 869 CTGGAGTGTGACCTCTGCGAAAGGTCAATTTGAGTCCAGAGACAGACCCCACTGAG 928
DB 870 CTGGAGTGTGACCTCTGCGAAAGGTCAATTTGAGTCCAGAGACAGACCCCACTGAG 929
QY 929 GGGCTGCGGAGCTCTAGGGGAGGTCTCCGCGCGGATGTGTTCAAGGACTGAGCCTG 988
DB 930 GGGCTGCGGAGCTCTAGGGGAGGTCTCCGCGCGGATGTGTTCAAGGACTGAGCCTG 989
QY 989 GGGAGCGCCACCAACCAACCACTGCT 1013
DB 990 GGGAGCGCCACCAACCAACCACTGCT 1014

0100

RESULT 4

US-09-764-877-2804

Sequence 2804, Application US/09764877

Parent No. US20020147140A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC005

CURRENT APPLICATION NUMBER: US/09/764,877

FILING DATE: 2001-01-17

Application data removed - refer to PALM or file wrapper

This Page Blank (uspto)

PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 546
LENGTH: 1288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (95) .. (928)
US-10-037-270-546

this date is good

Query Match 100.0%; Score 834; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGCTTGGGCGGAGCGGAGAGCTGGAGACCCCGCGCGGAGGCGGACCTCCGAC	60
DB	95	ATGCTTGGGCGGAGCGGAGAGCTGGAGACCCCGCGCGGAGGCGGACCTCCGAC	154
QY	61	CAGCGGCCCTTCTGTATAGGGGTGAGCGCGGCACTGCCAGCGGAACTGACCTGTGT	120
DB	155	CAGCGGCCCTTCTGTATAGGGGTGAGCGCGGCACTGCCAGCGGAACTGACCTGTGT	214
QY	121	GAGAAATCATGAGTGTCTGGGACGAAAGAGGTGAAACAGCGGACGGAAGGTGTCT	180
DB	215	GAGAAATCATGAGTGTCTGGGACGAAAGAGGTGAAACAGCGGACGGAAGGTGTCT	274
QY	181	ATCTGAGCGAGAGCAGTGTCTCAAGAGTCTGACGGCGAGAGAGGCGAAGGCTTG	240
DB	275	ATCTGAGCGAGAGCAGTGTCTCAAGAGTCTGACGGCGAGAGAGGCGAAGGCTTG	334
QY	241	AAAGGACAGTACATTTTTCAGATCCAGATGCTTGTATATATTTTATGACAGACT	300
DB	335	AAAGGACAGTACATTTTTCAGATCCAGATGCTTGTATATATTTTATGACAGACT	
QY	301	CTGAAGAAATCGTGGAGGGCAAAACGGTGGAGGTCCGACCTATGATTTTGTACAC	
DB	395	CTGAAGAAATCGTGGAGGGCAAAACGGTGGAGGTCCGACCTATGATTTTGTACAC	
QY	361	TCAGGTTACAGAGACCAAGGTGTCTACCTGCGGAGAGTGTCTGTAGGGGATC	420
DB	455	TCAGGTTACAGAGACCAAGGTGTCTACCTGCGGAGAGTGTCTGTAGGGGATC	514
QY	421	TTGCTGTCTAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTTCTTGTGACACC	480
DB	515	TTGCTGTCTAAGCCAGAGAGATCCGGGACATGTTCCACTGCGCTTCTTGTGACACC	574
QY	481	GACTCCGAGTCAAGGCTGTCTGAAGAGTTCCTCCGGAGAGTCCGCGGAGGAGGACTG	540
DB	575	GACTCCGAGTCAAGGCTGTCTGAAGAGTTCCTCCGGAGAGTCCGCGGAGGAGGACTG	634
QY	541	GAGCAGATTTCTGACGAGTACACCACTTCGTGAAGCGGCTTCGAGAGTTCGCTG	600
DB	635	GAGCAGATTTCTGACGAGTACACCACTTCGTGAAGCGGCTTCGAGAGTTCGCTG	694
QY	601	CCGACAAAGATATAGCGATGTGATCATCCAGAGAGTGCACATATGTGGCCATC	660
DB	695	CCGACAAAGATATAGCGATGTGATCATCCAGAGAGTGCACATATGTGGCCATC	754
QY	661	AACTGATGTCGAGACATCCAGGACATTCGAATGGGACATCGCAATGGGACCGA	720
DB	755	AACTGATGTCGAGACATCCAGGACATTCGAATGGGACATCGCAATGGGACCGA	814
QY	721	GAGGGTCCATGAGCGGAGCTTCAAGGCGACCTTTTGTAGCCAGGGAGCCACTGG	780
DB	815	GAGGGTCCATGAGCGGAGCTTCAAGGCGACCTTTTGTAGCCAGGGAGCCACTGG	874
QY	781	ATGCTGACCTTGGGAAAGGTCATATTTGAGTCCAGAGACAGCCCACTGA	834
DB	875	ATGCTGACCTTGGGAAAGGTCATATTTGAGTCCAGAGACAGCCCACTGA	928

US-10-037-270-546
Sequence 546, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungang
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725

This Page Blank (uspto)

APPLICANT: Dmanac, Radoje T.
 TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 FILE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/09/620,312D
 CURRENT FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1105
 SOFTWARE: pc_Fl_genes Version 1.0
 SEQ ID NO 546
 LENGTH: 1288
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (95)..(928)
 US-09-620-312D-546

Alignment Scores:
 Pred. No.: 9.34e-187 Length: 1288
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)

```

QY      1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB      95 ATGGCTTCGGCGGAGGCGGAAGACTGCGGAGCCCGCGGAGGCGGACCGTCCGAC 154
QY      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB      155 CAGCGGCCCTTCCCTGATAGGGGTGAGCGGCGGCACTCCAGCGGGAAGTGCAGCTGTGT 214
QY      41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 1
DB      215 GAGAGAGATCATGAGTGTCTGGACAGACGAGTGCACACCGCGCAGCGGAGGTGTC 1
QY      61 IleuSerGlnAspArgPheTyLysValLeuThrAlaGluGlnLysAlaLysVal 1
DB      275 ATCCTGAGCCAGACAGACAGTTCACAGTCTCTGACGCGCAGACGAGCCAGCTTG 334
QY      81 LysGlyGlnTyraenPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB      335 AAAGGACAGTACAAATTTGACATCCAGATGCCCTTTGATATGATTTGATGACAGGACT 394
QY      101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis 120
DB      395 CTGAAGACATCTGAGAGGCGAAGACGAGTGCAGACTTATGATTTGTGACACAC 454
QY      121 SerArgLeuProGluThrThrValValTyraProAlaAspValValLeuPheGluGlyTle 140
DB      455 TCAAGTTTACAGAGACACACGAGTGTCTACCTCGGACGCGTGTCTGTGAGGACAC 514
QY      141 LeuValPheTyraSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB      515 TTGGTGTCTACAGCAGACAGATCCGGACATGTTTCCACTGCGCTCTTGTGAGACAC 574
QY      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
DB      575 GACTCCGACGTCAGGCTCTCTCGAAGACTTCTCCGCGACGTCGCGGAGGAGGAGGAGCTG 634
QY      181 GluGlnIleLeuThrGlnTyraThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB      635 GAGCAGATTCTGACGCACTACACACCTTCTGTAAGCGGCGCTTCAAGAGTTCGCTGCTG 694
QY      201 ProThrLysLysTyraAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
  
```

RESULT 2
 US-09-620-312D-546
 Sequence 546, Application US/09620312D
 Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chengma
 APPLICANT: Asundi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wehrman, Tom
 APPLICANT: Xue, Aildong J.
 APPLICANT: Yang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yundong
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhiwei
 APPLICANT: John Tillinghast

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 08:11:31 ; Search time 74 Seconds

(without alignments)
1652.204 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGEDCSFAPBADRP.....HPGMLTSGKSHLSSSRPH 277

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09896522/rnucat.21112003.184355.5023/app.query.fasta_1.455
-DB=Issued Patents NA -OFMT=fasta -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=p2n -NCEM=ext -HEAPSIZE=500 -MINTEN=0 -MAXLEN=2000000000
-USER=US09896522.cgn1.1.56 @rnucat.21112003.184355.5023 -NCPV=6 -ICPV=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	834	4	US-09-536-647-1
2	1450	100.0	1288	4	US-09-620-312D-546
3	476	32.8	5687	4	US-09-221-017B-368
4	451	31.1	651	4	US-09-134-001C-781
5	434	29.9	633	4	US-09-107-532A-1370
6	412	28.4	1230025	4	US-09-198-452A-1
7	384	26.5	1830121	4	US-09-557-88A-1
8	384	26.5	1830121	4	US-09-643-990A-1
9	370.5	25.6	10011	4	US-08-961-527-76
10	351.5	24.2	580073	4	US-08-545-528D-1
11	156.5	10.8	4403765	3	US-09-103-840A-2
12	156.5	10.8	4411529	3	US-09-103-840A-1

C	13	137.5	9.5	1830121	4	US-09-557-884-1	Sequence 1, Appl1
C	14	137.5	9.5	1830121	4	US-09-643-990A-1	Sequence 1, Appl1
C	15	122	8.4	973	4	US-09-221-017B-368	Sequence 536, App
C	16	117	8.1	401	4	US-09-702-705-1476	Sequence 1476, App
C	17	117	8.1	401	4	US-09-736-457-1476	Sequence 1476, App
C	18	106.5	7.3	1230025	4	US-09-198-452A-1	Sequence 1, Appl1
C	19	101	7.0	272	4	US-09-313-294A-128	Sequence 128, App
C	20	99.5	6.9	627	4	US-09-134-001C-948	Sequence 948, App
C	21	94	6.5	1554	4	US-09-252-991A-5678	Sequence 5678, App
C	22	94	6.5	1623	4	US-09-252-991A-5648	Sequence 5648, App
C	23	94	6.5	2109	4	US-09-252-991A-5548	Sequence 5548, App
C	24	93.5	6.4	591	4	US-09-252-991A-9859	Sequence 9859, App
C	25	93.5	6.4	651	4	US-09-252-991A-9848	Sequence 9848, App
C	26	93	6.4	10240	4	US-08-961-527-327	Sequence 327, Appl1
C	27	92	6.3	226	4	US-09-313-294A-6207	Sequence 6207, App
C	28	92	6.3	1221	4	US-09-620-312D-162	Sequence 162, App
C	29	91.5	6.3	4128	3	US-08-793-624-2	Sequence 2, Appl1
C	30	91.5	6.3	4131	1	US-08-420-235B-2	Sequence 2, Appl1
C	31	91.5	6.3	4131	5	PCT-US95-10194-2	Sequence 2, Appl1
C	32	91.5	6.3	20710	1	US-08-420-235B-1	Sequence 1, Appl1
C	33	91.5	6.3	20710	3	US-08-793-624-1	Sequence 1, Appl1
C	34	91.5	6.3	20710	5	PCT-US95-10194-1	Sequence 1, Appl1
C	35	91.5	6.3	35100	2	US-08-770-379-18	Sequence 18, Appl1
C	36	91.5	6.3	35100	3	US-08-757-669A-18	Sequence 18, Appl1
C	37	91.5	6.3	35100	4	US-09-230-371A-18	Sequence 18, Appl1
C	38	91	6.3	2046	4	US-09-328-352-1540	Sequence 1540, App
C	39	90	6.2	858	4	US-09-334-938-12	Sequence 12, Appl1
C	40	89.5	6.2	798	4	US-09-252-991A-4797	Sequence 4797, App
C	41	89.5	6.2	4212	4	US-09-252-991A-4748	Sequence 4748, App
C	42	89.5	6.2	4242	4	US-09-252-991A-4697	Sequence 4697, App
C	43	88.5	6.1	9404	2	US-08-343-101A-1	Sequence 1, Appl1
C	44	88.5	6.1	9404	3	US-09-183-688-1	Sequence 1, Appl1
C	45	88.5	6.1	9404	4	US-09-519-489-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: **Ho, Yen Sen**
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1

Alignment Scores:

Pred. No.: 4.51e-187
Score: 1450.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Gaps: 0

US-09-896-522-2 (1-277) x US-09-536-647-1 (1-834)

QY 1 MetAlaSerAlaGlyCyluApCyGluSerProAlaProGluAlaApArgProHis 20
DB 1 ATGGCTTCGGCGAGGAGCGAAGACTGCGAGACCCCGCGCGAGGCGAGCGATCGCGAC 60

```

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 61 CAGGGGCGCTTCGATAGGGGTAGCGGCGCATCGACCGGGAAGTGCACGGTGT 120
QY 41 GluLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 121 GAGAGATCATGAGATGCTGCTGGACAGACAGAGGTGACACGCCGCGGACCGGAAGTGTGTC 180
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysValAlaLeu 80
Db 181 ATCTGAGCCAGACAGATTCCTCAAGGTCTGACGCGACAGAGCGCAAGGCGCCTTG 240
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 241 AAAGAGACAGTACATTTTTCATCCATCCAGATGCTTTTGAATGATTTGATGCACAGACT 300
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 301 CTGAGAAACATCGTGGAGGCGAAACGGTGGAGGTGCCGACCTGATTTTGTGACACAC 360
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 361 TCAAGGTTCACAGACACACCGTGTCTACCTCGCGACGTGTCTGTTGAGGCGCATC 420
QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 421 TTGGTGTCTACACCGAGGAGATCCGGACATGTTTCCACCTGCGCTTCTGAGGACACC 480
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCCTCGGACGTCGCCCGACGAGGAGGACCTG 540
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 541 GAGCAGATTCGACGCGATCACCACTTCGTGAAGCGCGCTTCGAGAGTTCTGCTG 600
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 601 CCGCAAGAGATGATGCCGATGATCATCCACAGAGAGTGAACATATAGTGTGCCATC 660
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
Db 661 AACGTGATCGTGCAGCATCATCAGACATCTGATAGTGTGACATCTGGAATGCGACCGA 720
QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
Db 721 GGAAGGTTCATAGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGCGGACCACTGCGG 780
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluLysSerSerArgProHis 277
Db 781 ATGCTGACCTCTGGCAACGGTCAATTTGAGTCCAGACAGACACCCAC 831

```

```

; APPLICANT: Dymanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-09-620-312D-546
Alignment Scores:
Pred. No.: 9,34e-187 Length: 1288
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 4
US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGAGCTTCGGCGGAGGCGAAGACTCGAAGGCCCGCGCGGACCGACCGTCCGAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGCGGCCCTTCCTGATAGGGGTGACCGCGGACCTCCAGCGGAGAGTGCACGGTGT 214
QY 41 GluLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 215 GAGAGATCATGAGATGCTGCGGACAGAAAGAGTGAACACCGCGGAGGAGTGTGTC 274
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysValAlaLeu 80
Db 275 ATCTGAGCCAGACAGATTCCTCAAGGTCTTACGCGACAGGACAGGCGCCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGACAGTACAAATTTTTCATCCAGATGCTTGTGATATGATTTGATGCACAGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 395 CTGAGAAACATCGTGGAGGCGAAACGGTGGAGGTGCCGACCTATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 455 TCAAGGTTCACAGACACACAGTGTCTACCTCGCGACGATGTTCTGTTGAGGCGATC 514
QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 515 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCCACTGGCCTCTTCGAGACACC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 575 GACTCCGACGTCAGGCTGTCTCGAAGATTCCTCGGACGTCGCGGAGGAGGAGGACTG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 635 GAGCAGATTCGACGAGTACACACACTTCGTGAACCGGCTTCGAGGAGTTCGCTG 694
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220

```

Db 695 CCGCAAGAGATGATGCGATGATCATCCAGAGAGTGGACATATGTTGCCATC 754
 Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleSTPbHisArg 240
 Db 755 AACCTGATCGTGGACGATCCAGCATCTTGATGTGATCACTGCAATGGCACCGA 814
 Qy 241 GlyGlySerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyAspHisProGly 260
 Db 815 GGAGGGTCCATGGGGGAGACCTCAAGCCGACCTTTTCTGAGCCAGGGGACCACTGGG 874
 Qy 261 MetLeuThrSerGlyIleAspSerHisLeuGluSerSerArgProHis 277
 Db 875 ATGCTGACCTGTGGCAACCGTCATCTTGAGTCCAGCAGACCCAC 925
 RESULT 3
 US-09-221-017B-368
 Sequence 368, Application US/09221017B
 Patent No. 6444799
 GENERAL INFORMATION:
 APPLICANT: Ross, Bruce C.
 TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 NUMBER OF SEQUENCES: 1120
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 PAGE MILL ROAD
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FASTSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/221,017B
 FILING DATE: 23-DEC-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PP1182
 FILING DATE: 31-DEC-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PP1546
 FILING DATE: 30-JAN-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PP2911
 FILING DATE: 09-APR-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU98/01023
 FILING DATE: 10-DEC-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Morroy, Gladys H
 REGISTRATION NUMBER: 32,430
 REFERENCE/DOCKET NUMBER: 27340-20021.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-813-5600
 TELEFAX: 650-494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 368:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5687 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULAR TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: UNKNOWN
 ORIGINAL SOURCE:
 ORGANISM: PORPHYROMONAS GINGIVALIS
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1...5687
 US-09-221-017B-368

Alignment Scores:
 Prod. No.: 2,276-53 Length: 5687
 Score: 476.00 Matches: 91
 Percent Similarity: 70.14% Conservative: 57
 Best Local Similarity: 43.13% Mismatches: 47
 Query Match: 32.83% Indels: 16
 DB: 4 Gaps: 4
 US-09-896-522-2 (1-277) x US-09-221-017B-368 (1-5687)
 Qy 25 LeuIleGlyValSerGlyIleThrAlaSerGlyIleSerThrValCysGluIleMet 44
 Db 4706 ATATATGAGATGACAGCGGAGAGTGTCTCGGCAAGATGACATGTGTAAGCTCTGT 4765
 Qy 45 GluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgIleValIleLeuSerGln 64
 Db 4766 GAGGCTTGTGTCGAGAGAGAT-----GTGCTGTACTCTGTCTAC 4804
 Qy 65 AspArgPheTyrIleVal-----LeuThrAlaGluGlnIleValAlaLeuIle 81
 Db 4805 GATTACTATTACAGGCCAATATACACCTCTCCCTTGAGGAAAGAAAG----- 4855
 Qy 82 GlyIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 101
 Db 4856 -----CTGACTAGACCATCCCATGCTTTCGACACGATATGTCTCAGGATATT 4909
 Qy 102 LysAsnIleValGluGlyIleThrValGluValProThrTyrAspPheValThrHisSer 121
 Db 4910 CTCTCTCTGAGGAGGCAAAAGATAGAGCGTCCGCTATATCTTCTAGAGCACAAT 4969
 Qy 122 ArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeu 141
 Db 4970 CGTTTGCAAGAAAGTAACCTGTCTGCTGCAAGATGATGATCTGATGAGAAATCTG 5029
 Qy 142 ValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 161
 Db 5030 ATATTGAAACAAAGAGTGGCGGATCTGATGATGTAAGTATTCCTGCATACCGAT 5089
 Qy 162 SerAspValArgLeuSerArgArgValLeuArgAspValArg--ArgGlyArgAspLeu 180
 Db 5090 GCGGATATTGATGGCGCGCGCTTGCGCGCATGTCCAGAAACGGGAGCAGCAATATAG 5149
 Qy 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluIlePheCysLeu 200
 Db 5150 GATCGGATTTGGCACAAATACCTTCAGTACGTTGGGCTTATGCACAGAGATTTTGTGAA 5209
 Qy 201 ProThrLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 5210 CATCCAGCGGTACGCCGATCTGATCATTCGGGAAGTGGATTCAATTGGTGGCGCTC 5269
 Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeu 231
 Db 5270 TCATCTCTTGTGCAAAATAATCCGATCGGTGATT 5302
 RESULT 4
 US-09-134-001C-781
 Sequence 781, Application US/09134001C
 Patent No. 6380370
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 PRIOR FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 781
 LENGTH: 651

```

1 TYPE: DNA
2 ORGANISM: Staphylococcus epidermidis
3 US-09-134-001C-781
4
5 Alignment Scores:
6 Pred. No.: 1,5e-51 Length: 651
7 Score: 451.00 Matches: 91
8 Percent Similarity: 63.68% Conservative: 44
9 Best Local Similarity: 42.92% Mismatches: 61
10 Query Match: 31.10% Indels: 16
11 DB: 4 Gaps: 4
12
13 US-09-896-522-2 (1-277) x US-09-134-001C-781 (1-651)
14
15
16 25 LeuilegylValserGlyGlyThralAserylYysSerThrValCysGluIlyIleMet 44
17 : :::::::::::::::::::::
18 43 ATTTATGGCATATGCGAGGGGTTCCAGGCTCTGGGAAACACAGTACTTAATGACATATG 102
19
20 45 GluIleuLeuGlyGlnAsnGluValGluGlnIArgGlnIArgIysValValIleLeuSerGln 64
21 : :: ::::
22 103 AAGAAATTTA-----GAGGTCAATAGTGCACCTTTAGCACAA 141
23
24 65 AsparGherylYlys-----ValIleuThrIArgIuGlnIlySAlaIleuIlys 81
25 : :::::::::::::::::::::
26 142 GATTATTACTATPAAAGATCAATCAATTTAATTGAAAGAACGGCTTGAAACA----- 195
27
28 82 GlyIuGlnYrAsnPhesPheSisProAspAlaPheAspAsnAluMetHisArgThrLeu 101
29 : :::::::::::::::::::::
30 196 -----AATTATGACCAACCCCATTTGCAATTTGATTAATGATTAATTCATTAATTTA 246
31
32 102 LysAsnIleValGluGlyIleThrValGluValProThrYrAspPheValThrHisSer 121
33 : :::::::::::::::::::::
34 247 AAGATTTAAGAAATGAAACACGTAAGGTACCTTAAGTACGATTATTTCAACATACA 306
35
36 122 ArgLeuProGluThrThrValValYrProIlaAspValIleuPheGluGlyIleu 141
37 : :::::::::::::::::::::
38 307 CGTGAATAAGAAACAATGTCATTGATGCCAAAAGATGTTATTATTCGTGAAGGTATCTTT 366
39
40 142 ValPheYrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 161
41 : :: ::::
42 367 GCTTTAGAAACAACACATTCAGACATGATGATGAAAAATTTATGTCATATCGAT 426
43
44 162 SerAspValArgLeuSerArgArgValIleuArgAspValArg---ArgGlyArgAspLeu 180
45 : :::::::::::::::::::::
46 427 GCTGATTTACGAATTTACGTAGGCTTCAAGAGATACTTAAAGAGCGGTGCTACATG 486
47
48 181 GluGlnIleLeuThrGlnIlyThrThrPheValIysProIlaPheGluGlnPheCysLeu 200
49 : :::::::::::::::::::::
50 487 GAATGAGTAATTAATCAATATCTTAATGTAAGACCTATGACATGACATTTATGAA 546
51
52 201 ProThrIlySlyTyralAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
53 : :::::::::::::::::::::
54 547 CTTACAAAAGACAGCTGACATCATATTCCTGAAGAGGACAGACGAATTAAGTTGCATA 606
55
56 221 AsnIleuIleValGlnHisIleGlnAspIleLeuAsn 232
57 : :::::::::::::::::::::
58 607 GATATTATGACTACTPAAATCCAGCATCTCGTAAC 642
59
60
61 RESULT 5
62 US-09-107-532A-1370
63 / Sequence 1370, Application US/09107532A
64 / Patent No. 6583275
65 /
66 / GENERAL INFORMATION:
67 / APPLICANT: Lynn A Doucette-Stamm
68 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
69 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
70 /
71 / NUMBER OF SEQUENCES: 7310
72 /
73 / CORRESPONDENCE ADDRESS:
74 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
75 / STREET: 100 Beaver Street
76 / CITY: Waltham
77 / STATE: Massachusetts
78 / COUNTRY: USA
79 /

```

```

1      ZIP: 02354
2      COMPUTER READABLE FORM:
3      MEDIUM TYPE: CD-ROM ISO9660
4      COMPUTER: PC
5      OPERATING SYSTEM: <unknown>
6      SOFTWARE: ASCII
7      CURRENT APPLICATION DATA:
8      APPLICATION NUMBER: US/09/107,532A
9      FILING DATE: 30-Jun-1998
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER: 60/085,598
12     FILING DATE: 14 May 1998
13     APPLICATION NUMBER: 60/051571
14     FILING DATE: July 2, 1997
15     ATTORNEY/AGENT INFORMATION:
16     NAME: Ariniello, Pamela Deneke
17     REGISTRATION NUMBER: 40,489
18     REFERENCE/DOCKET NUMBER: GTC-012
19     TELECOMMUNICATION INFORMATION:
20     TELEPHONE: (781)893-5007
21     TELEFAX: (781)893-8277
22     INFORMATION FOR SEQ ID NO: 1370:
23     SEQUENCE CHARACTERISTICS:
24     LENGTH: 633 base pairs
25     TYPE: nucleic acid
26     STRANDEDNESS: double
27     TOPOLOGY: circular
28     MOLECULE TYPE: DNA (genomic)
29     HYPOTHETICAL: NO
30     ANTI-SENSE: NO
31     ORIGINAL SOURCE:
32     ORGANISM: Enterococcus faecium
33     FEATURE:
34     NAME/KEY: misc_feature
35     LOCATION: (B) LOCATION 1...633
36     SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
37
38 US-09-107-532A-1370
39
40 Alignment Scores:
41 Pred. No.: 2,92e-49 Length: 633
42 Score: 434.00 Matches: 83
43 Percent Similarity: 62.33% Conservative: 51
44 Best local Similarity: 38.60% Mismatches: 65
45 Query Match: 29.93% Indels: 16
46 DB: 4 Gaps: 4
47
48 US-09-896-522-2 (1-277) x US-09-107-532A-1370 (1-633)
49
50 QY 22 ArgProPheLeuIleuGlyValSerGIyLThrAlaSerGIyLySerThrValCysGlu 41
51      :::::::::::::::::::::
52 Db 16 AAQCCAAATTAATTGGTGTGAAGCGGTGATCCGGAACGGAAAACAAGTGACGCCCT 75
53      :::::::::::::::::::::
54 QY 42 IyGIleMeGIuLeuLeuGlyIlnasngIuValGIugInArgInArgIyValValIle 61
55      |||
56 Db 76 GCAGTTTCAACAACACTCCCGGACCAT-----TCCATCATGATG 114
57      ::::
58 QY 62 LeuSerGIlnAspArgPheTYrIys-----ValLeuThraIaGIuGInIySaIaLYs 78
59      |||||
60 Db 115 CTGGAACAAGATCTTATATACAAAGATCAGAGCCATTTGAGTTTGAAGAACAACATTGAA 174
61      ::::
62 QY 79 AlaleuLyGIyGIlnTrAsnPhesphIsPrOkaspAlaPheAspAsnaRpleuMerHis 98
63      ::::
64 Db 175 ACA-----AATTACGATGATCACCCATTGGCTTTGATAATATGATCTGCTGATT 219
65      ::::
66 QY 99 ArgThrlleuLyAsnIleValGIuGIyLThrValGIuValProthrTYrAspPheVal 118
67      ::::
68 Db 220 CACAGATGTAGGCAGCTTACTGTAACTATAAGCATGTGAAAAACCGTGTATGATATGTA 279
69      ::::
70 QY 119 ThrIAserArgIeuProGIuThrThrValValIyTyProLaAspValIalleuPheGIu 138
71      |||
72 Db 280 GGCGATATCAAGAAGTAGGCAACGATCATPTCAGGAACCAAAGAAGTGATCATCTTGAA 339
73      ::::
74 QY 139 GIyIleuValPherIySerGIuGIuIleArGpserPheHlsleuArGleuPheVal 158

```


[illegible]


```
D6      146027 ATCATCGGCATGACAGGTGGCATCTGGATGTCAGCAGAAAGAATTCTAATCCGTTCCACTTGTTCA C   146089
OY          41 GtUySjIleMcGluleLeuglYglInasnGlValGIcUnlaRgclnArGlnARgLyVsAlVal    60
D6          146087 AAAGAGCTTGTGAATAAGTAAGCTGCCAACAAAATTTGGT----- 146125
OY          61 lileusSerGlnAspaRpPheTyLYvsVAlleuthrAlagluGlnlySalysAlaleu     80
Db          146126 ATCATTTACTGAAGACAGTTACTRTAAAGATCAAAGTCATTTAGAAATGACTGAGCGGGTA   146185
OY          81 lysglYglntyrASnpHeasphNIsproASPalaPhesApnaslpseumehNsArgThr    100
Db          146186 AAA----- ACGAATTAGCATGATCCAAACTGTATGGATCGCATTTACTTACCACAAT   146239
OY          101 LeulyASnlIlleValGIcUglyLytHthrValGIslvalProThrtYtyrspphealThrhIs    120
Db          146240 TTTTTAAAAATCTTAAAAAATGGCAGTGCAGTAGATGTGCTGTATTATAGCTTATGTAGAAACAT   146299
OY          121 SerArgyleProGLunthrrHthrValValTYrPrrolAspyAlaleupehiglUYtle     140
Db          146300 ACTCGACCAACAGAAACAACAGCATTTTCACACCACAAACGAATCGTAATTTTAAAGAGGATT   146355
OY          141 leuvAlpheTyserGlngllullearAspmecPhehisleuatrgyeupeheVAAspthr    160
Db          146360 TTATTACTTAACCGATGAACGAGTAGCCCAATTAGCCGATTTTCTGTGATTTCGRAGCAC   146419
OY          161 AspsEraSpvaIArgleUsErArGaTvallLeuarAspVaI--ArGarGlgYaTrgaSP    179
Db          146420 CCACCTGGATTATTTGTTTCATCCGCCGTTTACAAACGTGATVTGAMAGAACCGTGCGCTT   146479
OY          180 lengluglnlleLeuthrcGlnTYrHthrPHevalLyberolaAPheglUGlnUpheGs     199
Db          146480 CTACACATCAGTATGTGATCAATATCGTGCAACCGTCGTCCAATGTTCTTACAAATTTATT   146539
OY          200 leuproTHrylsEtyZylaaSpvAlllelePROArGilyAlaspasmeatValAla    219
Db          146540 GAGCCGCTCAAAACAARTCGGATATTGTCAITTCCTCGTGTGTGTAANAATCGCATTTGCG   146599
OY          220 lleasnleullavalgnHisilleglnasplleleuasN       232
Db          146600 ATCAATATGTGTAAAGCTCAAAATCCTTCATTTTATGAAAT   146638

RESULT 9
US-08-961-527-76/c
; Sequence 76, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesh
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PJ
```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-76

Alignment Scores:
Pred. No.:      1,25e+38      length:      10011
Score:          370.50        Matches:       69
Percent Similarity: 66.89%    Conservative:   30
Best Local Similarity: 46.62% Mismatches:       48
Query Match:     25.55%      Indels:         1
DB:              4           Gaps:            1

US-09-896-522-2 (1-277) x US-08-961-527-76 (1-10011)
QY      85 AsnPhasPAsPiProAAsPaLaPheAsPaenAspLeuMecHisArgThrLeuLySAsnIle 104
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      10000 AACTATGACCATCCTTTTGCCTTGATTACAGACTTCATGATCGACGAAATTAAAGCAATTG 9941
QY      105 ValGIuGLyLysThyValGIuValProThrTYRAspPheValThriSSerAlaLeuPro 124
      ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9940 TTGGCAGGGCCGTCCGGTGCACATCCGACCTACGACTATAACAGACATACACGAGTAGC 9881
QY      125 GIuThrThrValValTYr-ProLaAspValValLeuPheGIuGLyIleLeuValPheTYr 144
      ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9880 AAGACCTTATCGTCAGAACCTCAAGATGCTTTATGCTTGAGGATTTTGCTTGAGAG 9821
QY      145 SerGIuGIuIleArxAspMetPheHisLeuArxLeuPheValAspThraSPSerAspVal 164
      ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9820 GACAAGCGCTGCCGATTTGATGCATCAAGATTTTGTGTGATACGATGACGATGTG 9761
QY      165 ArgLeuSerArxArxValLeuArgAspVal--ArxArgLYArxAspLeuGIuGLuIle 183
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9760 CGCATTAATTCGTGCTATCAAGCGTGAATGAGAGACGTGGCCGTGACCTTGATAGCGTT 9701
QY      184 LeuThrGIuTYrThrThrPheValIysProAlaPheGIuGLuPheCySLeuProThrLYs 203
      ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9700 ATTAACAGTACTTAGAGTGTGTGTCAAACCAAGTACCACCAAGTTTATCGAGTCAACTAAG 9641
QY      204 LysTYrAlaAspValIleIleProArxGIyValAspAsmMetValAlaIleAsnLeuIle 223
      ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9640 CGTATGCTGAATATGCTCATTCCTCGAAGGGTTAGCAATACCTGCTATCGACCTGTG 9581
QY      224 ValGIuHisIleGIuAspIleLeu 231
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db      9580 ACGACCAAGATTGCAAGAATTGTTG 9557

RESULT 10
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmented
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193PI
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 580073
```


Qy 219 -----AlaIleasneuleuleValGlnHisileGlnAspIleLeuAsn 232
Db 1218746 GAGTAGAGCCCGCAGCCGGGCGACTCCCGGATATTCTCCATCCCTCGAAAC-----CGT 1218693
Qy 233 GlyAspIleCysLysTrpHisArgGlyLysSerAsnGlyArgSerTrpLysArgThrPhe 252
Db 1218692 GGTGTACCCCGCTCGGGCGCGGTAGCGGCGATACCGGTTGCGGTATCGGCGCAATTTC 1218633
Qy 253 SerGluProGlyAspHisProGlyMet 261
Db 1218632 AGCGGCAACGATCGACACAGGCGCTG 1218606
RESULT 12
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 3,91e-05 Length: 4411529
Score: 156.50 Matches: 70
Percent Similarity: 39.16% Conservatave: 51
Best Local Similarity: 22.65% Mismatches: 127
Query Match: 10.79% Indels: 61
Gaps: 13
US-09-896-522-2 (1-277) x US-09-103-840A-1 (1-4411529)
Qy 3 SerIleAglyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHisGlnArg 22
Db 1219971 GCGGCCACCGCGGAATTCTTCGGCGAGCCCAACAAACCGGACCGCGC-----GTG 1219918
Qy 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLys 42
Db 1219917 CCGTTTCATCGGTGTGTGCGGCGAGTGTGGCGGTGCGGCAAGTCACACCGCGCGGTG 1219858
Qy 43 IleMetGluLeuLeuGlyGlnAsnGlnValGluGlnArgGlnArgLysValIleLeu 62
Db 1219857 CTCGAGGGCGCTGCTGCTGCTGGAGTACACACCGCGG-----GTAGATCTGGTG 1219807
Qy 63 SerGlnAspArgPheTrpLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGly 82
Db 1219806 ACCACCGACGGCTTTCTC---TACCCCAACCGCGAGCTGCGACGCGCAACCTCATGCAT 1219750
Qy 83 GlnTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
Db 1219749 CGCAAGGTTT-----CGGAAAGCATATACCGCGCGGCGACATGATCGGTTTGTACAC 1219696
Qy 103 AsnIleValGluGlyLysThrValGlu---ValProThrTrpAspPheValThrHisSer 121
Db 1219695 TCGGTGAAGTCGCGTCCGACTAGCATGTGCGCGGTGTATTCGCAATTGCACTAGAC 1219636
Qy 122 ArgLeuPro-----GluThrThrValValTrpAlaAspValValLeuPheGluGly 139
Db 1219635 ATCATCCGGGGGAGCAACAGTGTCTCGCATCC---GACATCTGATCTCGAGGGG 1219579

Qy 140 IleLeuValPheTrpSerGlnGlu-----IleArgAspMetPheHisLeuArgLeu 156
Db 1219578 CTCAGCTCTTGAGAGACCGGCGCCAGCTCATGTGTCTCGATCTGTCATTTTGCTTC 1219519
Qy 157 PheValAspThrAsp---SerAspValArg-----LeuSerArgArgValLeuArg 172
Db 1219518 TATGTAGACCGCCGATCGAGCATCGAGACAGATGTAGTATCAGCGTTTGTGGCCANG 1219459
Qy 173 AspValArgArgGlyArgAspLeuGlnIleLeuThrGlnThrThrPheVal--- 191
Db 1219458 CGCACACCGCGTTTCCCGCCGACCGGATACACTTCCACCACTAGCGGCTTTTCTGAC 1219399
Qy 192 -----LysProAlaPheGlu 196
Db 1219398 TCGCAGGCGCTGCTCGCGCGCGCGAGATCTGCGGACCATCAACCGGCCAACCTCTCG 1219339
Qy 197 GluPheCysLeuProThrLysLysTrpAlaAspValIleIleProArgGlyValAspAsn 216
Db 1219338 GAGAACATTCTGCGACCGCTCCCGGCGCACCTGTGCTGCGCAAGACGCGCATCAT 1219279
Qy 217 MetVal----- 218
Db 1219278 TCCATTACCGGCTGCGGCTGCGCAAGCTGTAATTGGCGGTCCATTCGCCGGTGCAGGC 1219219
Qy 219 -----AlaIleasneuleuleValGlnHisileGlnAspIleLeuAsn 232
Db 1219218 GAGTAGAGCCCGCAGCCGGGCGACTCCCGGATTTCCATCCCTCGAAAC-----CGT 1219165
Qy 233 GlyAspIleCysLysTrpHisArgGlyLysSerAsnGlyArgSerTrpLysArgThrPhe 252
Db 1219164 GGTGTACCCGCTCGGGCGCGGTAGCGGCGATACCGGTTGCGGTATCGGCGCAATTTC 1219105
Qy 253 SerGluProGlyAspHisProGlyMet 261
Db 1219104 AGCGGCAACGATCGACACAGGCGCTG 1219078
RESULT 13
US-09-557-884-1/c
Sequence 1, Application US/09557884
Patent No. 650581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

```

```

Alignment Scores:
Pred. No.: 0.00341 Length: 1830121
Score: 137.50 Matches: 52
Percent Similarity: 41.70% Conservative: 46
Best Local Similarity: 22.13% Mismatches: 86
Query Match: 9.48% Indels: 51
DB: 4 Gaps: 7

```

```

US-09-896-522-2 (1-277) x US-09-557-884-1 (1-1830121)
QY 23 Prothetleuilegylvalserglylthralaserglylserthralcyeglyllys 42
DB 670207 CCTTACATTAATTAATGCTGCGACAGTGTGCTGTTGGGAAAAGCACTTTCGACGTAAT 670148
QY 43 Ilemergluleuenglyglnasngluvalgluglnarglnarglyvalleu 62
DB 670147 CTACATCTCTACTTCTCATTTGCCAATCGAA-----AGAAAAGTTGATCTCATC 670097
QY 63 SerGlnaspargphetyrlysvallleuthralaeglulnlysalaleuylgly 82
DB 670096 ACAACAGATGGGTTCTTTATCCCTTAACAAAGCTTAAACAAGTAATCTTTTAACAAAAG 670037
QY 83 GlnTyAsnPhaspHisProaspAlaPheaspAsnaspLeuMetHisArgThrLeuLys 102
DB 670036 AAA-----GGTTTTCCCGTTTCTTATGACACACCTAAACCTATTCGCTTTTACCG 669986
QY 103 AsnIleValgluglylsth--ValgluValProThTyAspPheValThriSer 121
DB 669985 GATGTAAATCAAGGTAAAGCAATGTTACAGCACTTAATTAATCTCATTAACATACGAT 669926
QY 122 ArgLeuProgluthr--ThrValValTyProAlaAspValValLeuPhegluglyl 140
DB 669925 ATTATTTCCGATTAATTTGATGTGAGATTAACCTCAATTTCTTAATTTAGAGGCTTG 669866
QY 141 LeuValPheTySerSerIngluileArg--AspMetPheHisLeuArgLeuPheValAsp 159
DB 669865 AATGTACTTCAACGGGTAATTAATAAACCGATCAACATTCGATTAAGATTITGCGAT 669806
QY 160 ThrAspSeraspValArgLeuSerArgValLeuArgAspValArgArglyArgAsp 179
DB 669805 TTCTCTATTATGTTGATGCGAAGAAAATTATTTGAAA----- 669767
QY 180 LeuGlulnleuThrGlnTyThrThrPheValLysProAlaPhe----- 195
DB 669766 ---GAATGATATATCAACAGATTTCTGAAATTCGCGAAAGTGATCAACGATCTTAAT 669710
QY 195 ----- 195
DB 669709 TCTTACTTTAAACATTACGCAAGCTTATCAAAAGAAAGGCTATCGGACGCGCAAGTAAA 669650
QY 196 -----GluGluPheCysLeuProThTyLysTy 205
DB 669649 ATCTGGAGTAATAATTAACGATTAATCTCAATCAAAAACATTTCTTCACCCGGAACGA 669590
QY 206 AlaAspVallelleProArglyValAspAsnMetVallelle 220
DB 669589 GCAAAATTAAATTTAAAAAAGGTCAATATCAACCAAGTAAATTA 669545

```

```

RESULT 14
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams

```

```

; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS V6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186PIC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1
;
; Alignment Scores:
; Pred. No.: 0.00341 Length: 1830121
; Score: 137.50 Matches: 52
; Percent Similarity: 41.70% Conservative: 46
; Best Local Similarity: 22.13% Mismatches: 86
; Query Match: 9.48% Indels: 51
; DB: 4 Gaps: 7
;
US-09-896-522-2 (1-277) x US-09-643-990A-1 (1-1830121)
QY 23 Prothetleuilegylvalserglylthralaserglylserthralcyeglyllys 42
DB 670207 CCTTACATTAATTAATGCTGCGACAGTGTGCTTGGGAAAAGCACTTTCGACGTAAT 670148
QY 43 Ilemergluleuenglyglnasngluvalgluglnarglnarglyvalleu 62
DB 670147 CTACATCTCTACTTCTCATTTGCCAATCGAA-----AGAAAAGTTGATCTCATC 670097
QY 63 SerGlnaspargphetyrlysvallleuthralaeglulnlysalaleuylgly 82
DB 670096 ACAACAGATGGGTTCTTTATCCCTTAACAAAGCTTAAACAAGTAATCTTTTAACAAAAG 670037
QY 83 GlnTyAsnPhaspHisProaspAlaPheaspAsnaspLeuMetHisArgThrLeuLys 102
DB 670036 AAA-----GGTTTTCCCGTTTCTTATGACACACCTAAACCTATTCGCTTTTACCG 669986
QY 103 AsnIleValgluglylsth--ValgluValProThTyAspPheValThriSer 121
DB 669985 GATGTAAATCAAGGTAAAGCAATGTTACAGCACTTAATTAATCTCATTAACATACGAT 669926

```

```

Qy 122 ArgLeuProGluThr---ThrValValTyrProAlaAspValLeuPheGluGlyIle 140
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 669925 ATATATCCCGATAAATTGATGCTGTAGATTAACCTGACATCTTATTATTTGAAGCGCTTG 669866
Qy 141 LeuValPheTyrSerGlnGluIleArg---AspMetPheHisLeuArgLeuPheValAsp 159
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 669865 AATGTACTTCAACGCGGTATATATATAAACCGATCAACATTCGTATCAGATTTTGTGGAT 669806
Qy 160 ThrAspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAsp 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 669805 TTCTCTATTATATGTTGATGCTGAGAAAGAAATATTATTTGAAA----- 669767
Qy 180 LeuGluGlnIleLeuThrGlnTyrThrThrPheValIleProAlaPhe----- 195
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 669766 ---GAATCGTATATCAACGATTTCTGAAATTCGCGCAAAATGCATTCACAGATCTTAAT 669710
Qy 195 ----- 195
Db 669709 TCTTACTTTAAACATTACGCAAGCTTATCAAAAGAGAGCGCTATCGCGACGCAAGTAAA 669650
Qy 196 -----GluGluPheCysLeuProThrIleValGlyTyr 205
Db 669649 ATCTGGGATGAATTACGAGATTAAATCTCAATCAAAACATTCCTCCAAACCCGCGAGCGA 669590
Qy 206 AlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 669589 GCAATTTAATTCTTAAABAAAGATCATATATACCAAGTGAATTA 669545

RESULT 15
US-09-221-017B-536
; Sequence 536, Application US/09221017B
; Patent No. 644799
;
; GENERAL INFORMATION:
; APPLICANT: ROSE, BRUCE C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221.017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141

```

```

; INFORMATION FOR SEQ ID NO: 536:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...973
; US-09-221-017B-536

Alignment Scores:
Pred. No.: 1,44e-06 Length: 973
Score: 122.00 Matches: 35
Percent Similarity: 46.21% Conservative: 26
Best Local Similarity: 26.52% Mismatches: 51
Query Match: 8.41% Indels: 20
DB: 4 Gaps: 4

US-09-896-522-2 (1-277) x US-09-221-017B-536 (1-973)
Qy 62 LeuSerGlnAspArgPheTyrIleValLeuThrAlaGluGlnIleValAlaValAlaLeuLys 81
Db 27 CTTTCGCTGATATACATTTTC-----ATCAACCGGAGAGATTCCGACCGCATGATCC 80
Qy 82 GlyGlnTyrAsnDheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 101
Db 81 GGAGACTATGATTTTCGAATCGCTCTACGCTTGTGATCTGCCCTTTTCATATAAGATCTG 140
Qy 102 LysAsnIleValGluGlyIleValGluValProThrTyrAspPheValThrIleSer 121
Db 141 AAGCAGATGATCCCGGTGAAGAGAGAGCTGCTCACTATGACTTCGCAACCGGAGATG 200
Qy 122 ArgLeu-----ProGluThrThrValValTyrProAlaAspValLeuPheGluGly 139
Db 201 AGGCTCTACAAGAAATACCATACAGCTGAAGAAGAGAGATATCTGTATCTCCGAAGCG 260
Qy 140 IleLeuValPheTyrSerGln-----GluIleArgAspMetPheHisLeuArgLeu 156
Db 261 ATTCATGCGCTGATCCGAGCTGATACCCGCGTGCCTCCGAATCAAGTACGTATAAGATC 320
Qy 157 PheValAspThr-----AspSerAsp 163
Db 321 TATGTAGCGCTCTCACTGCATATGATTTGATGCTCACATCGATCCCGACGTACGGAC 380
Qy 164 ValArgLeuSerArgArgValLeuArgAspValArg 175
Db 381 AACCGACTATCAAGCGGATTTGTGCGCGACTATCGC 416

```

Search completed: November 25, 2003, 12:12:34
 Job time : 7528 secs


```
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 361 TCAGAGTTTACAGACGACCGTGTCTACCCGCGAGCGTGTCTGTGTTGAGGGCATC 420
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 421 TTGGTGTCTTCAACCGACGAGATCCGGGACATGTTCCACCTCGGCCCTTCTGTGAGACCC 480
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCCGAGGGAGGACCTG 540
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIlePheGluGluPheCysLeu 200
Db 541 GAGGAGATTCGACGACGATACCACTTGTGTAAGCCGGCTTCGAGAGGTTCTGCTG 600
QY 201 ProThrIleValTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 601 CCGCAAAAGAGTATGCGATGTATCATCCACGAGAGTGGACATATGTTGCCATC 660
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheArg 240
Db 661 AACCTGATTCGTCAGCATCAGAGACATTCGATGATGATGATGATGATGATGATGATG 720
QY 241 GlyIleSerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyIleAspHisProGly 260
Db 721 GGAAGGTTCATGAGCGAGATACAGAGGACCTTTTCTGAGCGAGGAGACACCTTGGG 780
QY 261 MetLeuThrSerGlyIleArgSerHisLeuGluSerSerSerArgProHis 277
Db 781 ATGTGACCTCTGGCAAAAGCTCACATTTGGAGTCCAGACAGACACCCAC 831

RESULT 2
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Weinman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunru
; APPLICANT: Wang, Zhimei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc FL_genes Version 1.0
; SEQ ID NO 546_
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
```

```
US-10-037-270-546
Alignment Scores:
Pred. No.: 7,99e-287 Length: 1288
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-896-522-2 (1-277) x US-10-037-270-546 (1-1288)
QY 1 MetAlaSerAlaGlyIleGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACGTCGCCAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyIleSerThrValCys 40
Db 155 CAGCGCCCTTCTCTGATAGGGGTGAGCGCGGACCTGCGACGCGGAGTGCACCGTGT 214
QY 41 GluIleIleMetGluLeuLeuGluGlnAsnGluValGluGlnIleAsnArgIleVal 60
Db 215 GAGAGATATCATGAGTTGCTGGGACAGAACAGAGTGCAGAGCGGAGCGAGAGGTGTC 274
QY 61 IleLeuSerGlnAspArgPheTyrIleValIleThrAlaGluGlnIleValAlaVal 80
Db 275 ATCTTACGACGAGACAGTTCTTCAAGGTCCTGACCGGACGAGAGGAGGAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGCAAGTACATTTTGCATCCATCCAGATGCTTGTATATGATTTGATGACAGACT 394
QY 101 LeuIleAsnIleValGluGlyIleThrValGluValProThrTyrAspPheValThrHis 120
Db 395 CTAAAGATCATCTGAGAGGCGCAAAACGTGAGAGTCCGACCTATATTTGTGACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 455 TCAGGTTACAGACGACGAGTGTCTACCCGCGGACGTTGTTGAGGACATC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 515 TTGGTGTCTTCAACGACGAGAGATCCGGGACATGTTCCACTGCGCTTCTGTGACACC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGAGGACCTG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIlePheGluGluPheCysLeu 200
Db 635 GAGCAGATTCGACGACGATACCACTTCTGTGAACCCGGCTTCGAGGAGTTCTGCCCTG 694
QY 201 ProThrIleValTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 695 CCGACAAAGAGTATCCGATGTGATCATCCGCGAGGTGACAAATGTGTTGCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheArg 240
Db 755 AACCTGATTCGTCAGCATCAGACGATTCGTAAGTGACATTCGCAATGACACCGA 814
QY 241 GlyIleSerAsnGlyArgSerTyrIleArgThrPheSerGluProGlyIleAspHisProGly 260
Db 815 GAGAGGTCCATGCGGAGGAGCTCAAGCGGACCTTTTCTGAGCAGAGGAGGACCCCTGGG 874
QY 261 MetLeuThrSerGlyIleArgSerHisLeuGluSerSerSerArgProHis 277
Db 875 ATGTGACCTCTGGCAAAAGCTCACATTTGGAGTCCAGACAGACACCCAC 925

RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
```

? APPLICANT: Glucksmann, Maria A.
 ? TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
 ? TITLE OF INVENTION: US\$ THEREOF
 ? FILE REFERENCE: 381552001700
 ? CURRENT APPLICATION NUMBER: US/09/896,522
 ? PRIOR FILING DATE: 2001-06-28
 ? PRIOR APPLICATION NUMBER: 60/216, 503
 ? PRIOR FILING DATE: 2000-06-30
 ? NUMBER OF SEQ ID NOS: 6
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 1
 ? LENGTH: 1624
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: (94)...(927)
 ? OS-09-896-522-1

Alignment Scores:

Pred. No.:	9, 84e-287	Length:	1632
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-2 (1-277) x US-09-896-522-1 (1-1624)

QY	1	MetAlaSerAla61G1yGluIuAspProGluSerProAlaProGluAlaAspArgProHis	20
Db	94	AlRGCTTCGGCGGGAGCGAGACCTGGAGACCCCGCGCGAGGCCGACCGTCGGCAC	153
QY	21	GlnArgProPheLeuIleGlyValSerGlyGlyThralaSerGlyLysSerThrValCys	40
Db	154	CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACCTGCCAGCGGGAAAGTCACCGTGTGT	213
QY	41	GluLysIleMetGluLeuLeuGlyGlnMetGlnValGluGlnArgGlnArgLysValVal	60
Db	214	GAGAAATATGAGAGCTGCTGGGACAGAACAGAGGTGMAACGGCGCGGAGAGGTGCTC	273
QY	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrIaGluGlnLysValaLysAlaLeu	80
Db	274	ATCCTGAGCGCAGACAGAGTCTCAAGAGTCTTGACGCGAGAGCAGAAAGCCAGGCTTGC	333
QY	81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr	100
Db	334	AAAGGACAGTAAATTTGACCATCCAGATGCTTGTATATGATTGATGACACAGAGCT	393
QY	101	LeuLysAsnIleValGluGlyLysThrValGlnValProThrTyrAspPheValThrHis	120
Db	394	CTGAAGAAATGTTGAGAGGCAAAACGGTGGAGGTGGCGCATATGATTTGGTGCACAC	453
QY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe	140
Db	454	TCAAGGTTACAGAACACACACGCTGCTCAACCTCGACAGCGGTCTCTGTTGAGGCGATC	513
QY	141	LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr	160
Db	514	TTGGTGTTCTTACAGCAGGAGATCCGGAGACATGTTTCCACCTGCCTTCCTGGTGCACCC	573
QY	161	AspSerAspValArgLeuSerArgArgValLeuArgAspValArgAspArgLysArgAspLeu	180
Db	574	GACTCCGAGCTAGGCTGTCTGAAAGAGTCTCCGGAGCGGCGCCGAGGAGAGGACCTG	633
QY	181	GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu	200
Db	634	GAGCAGATTCTGACCAAGTACACCACTTCGTGAAAGCGGCGCTTCGAGGAGTTCGTGCTG	693
QY	201	ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
Db	694	CCGACAAAGAAATATGCGGATGTATCATCCACAGAGAGTGCACATATAGTGTTCATC	753

Qy	Db	Qy	Db
222	754	241	814
AsnLeuLeuValGlnHisIleLeuAspIleLeuLysAspIleCysValSTPRHisArg	AACTATATGTTGAGACACATCCAGGACATTCCTGAAATGTATGACATCTGCATAATTTGGACCGA	GLGILSerAsnGILValSerSerTyrIleYsaRqTHrPheSerGILupProGILYAspHisProGILY	GGAGGGGTCMAATGGGGCGAGCTTACCAAGGGAGACCTTTTGTAGGCCAGGGGGACCAACCTTGGG
261	874	261	874
MetLeuThrSerGILYLYsaRSerHisIleuGlnIuSerSerSerArpProHis	ATCTTAACCTCTGGCAAGGTCACATTTTGGAGTCTACGACACAGACCCAC		924

RESULT 4

```

US-09-833-381-2048
/ Sequence 2048, Application US/09833381
/ Patent No. US20020132090A1
/ GENERAL INFORMATION:
/ APPLICANT: Robison, Keith E.
/ TITLE OF INVENTION: No. US20020132090A1 Nuclear Acid and Protein Homologs
/ FILE REFERENCE: 5800-119
/ CURRENT APPLICATION NUMBER: US/09/833,381
/ CURRENT FILING DATE: 2001-04-11
/ PRIOR APPLICATION NUMBER: 09/516,448
/ PRIOR FILING DATE: 2000-02-29
/ NUMBER OF SEQ ID NOS: 2050
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2048
/ LENGTH: 1648
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(1648)
/ OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

```

Alignment Scores:

Pred. No.:	8, 44e-273	Length:	164e
Score:	264.00	Matches:	277
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	0
Query Match:	95.31%	Indels:	1
DB:	10	Gaps:	0

US-09-896-522-2 (1-277) X US-09-833-381-2048 (1-1648)

QY	1	MetAlaSerAlaG1G1G1G1nAspProY6G1uSerProAlaProGluAlaAspArgProHis	20
Db	84	ATGGCTCCGGCGGAGCGAGAGACTGGACAGAGCCCCGGCGCGAGCGCAACGTCGCCAC	143
QY	21	GlnArgProPheLeuI16G1ValSerG1G1Thra1AsaSerG1Y1ySerThrValCys	40
Db	144	CACGGGCCCTTCCTCTGATAGGGGTAGCGCGCGGCATCGCACGGGGAAGTCGACCGTGTCT	203
QY	41	GluY1eI1MeerGluLeuLeuG1G1nAsnGluValG1uGlnArgG1nArgY1eVal1	60
Db	204	GAGAAATCATGAGATGGCTGCGGACAGAACAGAGGTGGAAACAGCGGACGGAAAGTGTCT	263
QY	61	I1eIeuSerG1nAspArgPheYr1y1yVal1LeuThr1aG1uG1nY1eVal1aY1eAla1eU	80
Db	264	ATCTTGAGCGAGACAGGTCTTACAAAGTCTCTGACGCGACAGCGAAGGGCCAAAGCCCTTG	323
QY	81	Y1eG1G1n1yAsn1PheAspH1e1ProAsp1aPheAspAsnAsp1eUmeTh1aArgThr	100
Db	324	AAAGGACATACAAATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGACACAGACT	383
QY	101	LeuY1eAsn1eValG1uG1Y1y1ySerThrValG1uVal1ProThr1yAspPheVal1ThrHis	120
Db	384	CTGAGAACATCGTGGAGGGCAAAACGGGTGAGGGCGCACTATGATTTTGTGACACAC	443
QY	121	SerArgLeuProG1nThrThrValVal1yProAlaAspValVal1eU1ePheG1uG1y1e	140
Db	444	TCAGAGTTTACCGAAGCACCGTGTGTCAACCTCGACGCGGTTCGTTGTTGAGGGCATC	503

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 504 TTGGTGTCTTACACCGACGAGATCCGGACATGTTCCACCTGCGCTCTTCTGTGACACC 563
QY 161 AspSerAspValaLysLeuSerArgArgValaLeuArgAspValaArgArgIleArgAspLeu 180
DB 564 GACTCCGACGTACAGCTGTCTCGAAGATTCTCCGGAGCTGGCCGACGAGGAGGAGCTTG 623
QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValaLysProAlaPheGlnGluPheCysLeu 200
DB 624 GAGCAGATTCTGACGACGATACACCACTTGTGTAAGCGGCTTCGAGGAGTTCTGCTG 683
QY 201 ProThrIleLysTyrAlaAspValaIleLeuProArgIleValaAspMetValaAla 220
DB 684 CCGCAAGAAGATGTTCCGATGATCATCCCGAGGATGACATATGTTGCCAT 743
QY 220 eAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheSar 240
DB 744 CAACCTGATCTGACGACATCCGACATTTCTGAATGGTGACATCTGCAATGGCACCG 803
QY 240 gGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
DB 804 AGGAGGATCCATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCGAGGAGCACCTG 863
QY 260 yMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
DB 864 GATGCTGACCTTGCGCAACCGGTCACTTGAGTCCAGCAGCAGACCCAC 915

RESULT 5

US-09-918-995-30379
Sequence 30379, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30379
LENGTH: 472
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(472)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Alignment Scores:

Pred. No.: 5,16e-134 Length: 472
Score: 134.00 Matches: 134
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.38% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

QY 33 AlaSerGlyLysSerThrValaCysGlyLysIleMetGluLeuGluGlnAsnGluVal 52
DB 41 GCCAGCGGGAAGTGAACCTGTGTGAAGAATCATGAGTTGCTGGGACGAAGAAGTG 100
QY 53 GlnGlnArgGlnaGlyLysValaValaIleLeuSerGlnAspArgPheTyrLysValaLeuThr 72
DB 101 GAACAGCGGCGCAGGAGGTGTCTGAGCCAGCAGAGTTCTCAAGAGTCTTGACG 160
QY 73 AlaGlnGlnLysAlaLysAlaLeuLysGlyLysIleThrIleAsnPheAspAlaPhe 92

DB 161 GCAGACGAGAAGGCCAAGGCTTTGAAAGACAGTACAAATTTTGACCATCCAGATGCTTT 220
QY 93 AspAsnAspLeuMetHisArgThrLeuLysAsnIleValaGlnGlyLysThrValaGluVal 112
DB 221 GATPAGATATTGATGACAGACTCTGAAGAATCATGCTGAGAGGCAAAACGGGTGAGTG 280
QY 113 ProThrTyrAspPheValaThrHisSerArgLeuProGluThrThrValaTyrProAla 132
DB 281 CCACCTATGATTGTTGTGACACCTCAAGGTACCAAGACCAAGGATGATCTTACCTG 340
QY 133 AspValaIleLeuPheGlnGlyIleLeuValaPheTyrSerGlnGluIleArgAspMetPhe 152
DB 341 GACGTGCTTCTGTTAAGGACATCTTGCTGTTCTAAGCCAGAGATCCGGACATGTT 400
QY 153 HisLeuArgLeuPheValaAspThrAspSerAspValaArgLeu 166
DB 401 CACCTGCGCTTCTGTTGACACCGACTCCGACGTGAGGCTG 442

RESULT 6

US-10-029-386-15883
Sequence 15883, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029, 386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15883
LENGTH: 187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
OTHER INFORMATION: NT HIT: g114783235, EVALUE 1.00e-100
US-10-029-386-15883

Alignment Scores:

Pred. No.: 4,27e-55 Length: 187
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-10-029-386-15883 (1-187)

QY 218 ValaIleAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
DB 3 GTTGCCATCAACTGATCGTGACGACATCCAGACATCTGAATGATGTCATCTCCAA 62
QY 238 ThrHisArgGlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 257
DB 63 TGGCACCAGAGAGGCTCCATGGCGGAGCTTAACACCGACCTTTCTGAGCCAGGAGAC 122
QY 258 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277

```
Db      123  CACCTGGATGCTGACCTCTGGCAACGGTCACTTTGGATGATCCAGACAGACCCAC 182
RESULT 7
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: Bf64526.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: Bf64526.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183

Alignment Scores:
Pred. No.:      1,05e-54      Length:      510
Score:          60.00         Matches:      60
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    21.66%       Indels:      0
DB:             12           Gaps:       0

US-09-896-522-2 (1-277) x US-10-029-386-2183 (1-510)

QY      218  ValAlaIleasnleuileValGlnHisIleGlnAspIleuansnglyAspIleCysLys 237
Db      26  GTTGCCATCAACCTGATCGTGAGCAGCATCCAGACATTTGATGATGATCGCAAA 85
QY      238  TrpHisArgGlyGlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAsp 257
Db      86  TGGCACCAGAGAGGCTCCAAATGGCGGAGCTACAAAGCGAGCCTTTCTGAGCCAGGGGAC 145
QY      258  HisProGlyMetLeuThrSerGlyIlyArgSerHisLeuGlnSerSerSerProHis 277
Db      146  CACCTGGATGCTGACCTCTGGCAACGGTCACTTTGGATGATCCAGACAGACCCAC 205

RESULT 8
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804

Alignment Scores:
Pred. No.:      1.49e-53      Length:      9732
Score:          60.00         Matches:      60
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    21.66%       Indels:      0
DB:             10           Gaps:       0

US-09-896-522-2 (1-277) x US-09-764-877-2804 (1-9732)

QY      218  ValAlaIleasnleuileValGlnHisIleGlnAspIleuansnglyAspIleCysLys 237
Db      3901 GTTGCCATCAACCTGATCGTGAGCAGCATCCAGACATTTGATGATGATCGCAAA 3960
QY      238  TrpHisArgGlyGlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAsp 257
Db      3961 TGGCACCAGAGAGGCTCCAAATGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGGGAC 4020
QY      258  HisProGlyMetLeuThrSerGlyIlyArgSerHisLeuGlnSerSerSerProHis 277
Db      4021 CACCTGGATGCTGACCTCTGGCAACGGTCACTTTGGATGATCCAGACAGACCCAC 4080

RESULT 9
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2805

Alignment Scores:
Pred. No.:      2.75e-53      Length:      19125
Score:          60.00         Matches:      60
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    21.66%       Indels:      0
DB:             10           Gaps:       0

US-09-896-522-2 (1-277) x US-09-764-877-2805 (1-19125)

QY      218  ValAlaIleasnleuileValGlnHisIleGlnAspIleuansnglyAspIleCysLys 237
Db      5997 GTTGCCATCAACCTGATCGTGAGCAGCATCCAGACATTTGATGATGATCGCAAA 6056
QY      238  TrpHisArgGlyGlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAsp 257
Db      6057 TGGCACCAGAGAGGCTCCAAATGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGGGAC 6116
QY      258  HisProGlyMetLeuThrSerGlyIlyArgSerHisLeuGlnSerSerSerProHis 277
Db      6117 CACCTGGATGCTGACCTCTGGCAACGGTCACTTTGGATGATCCAGACAGACCCAC 6176

RESULT 10
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
```

```

; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-220

Alignment Scores:
Pred. No.: 7.73e-17 Length: 1310
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x US-09-925-300-220 (1-1310)

Qy 189 ThpHeVallyspProAlaPheGluGluPheCysLeuProThrIysTyrAlaAspVal 208
Db 789 ACGTTCGTCAGCCTGCTTGGAGAAATCTGCTTGCCACAAAGAAAGTATGCTGATGTG 848

Qy 209 IleIleProArgGly 213
Db 849 ATCATCCCTAGAGGT 863

RESULT 11
US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098, 841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (230) ..(1075)
US-10-098-841-53

Alignment Scores:
Pred. No.: 8.21e-17 Length: 1402
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 13 Gaps: 0

US-09-896-522-2 (1-277) x US-10-098-841-53 (1-1402)

Qy 189 ThpHeVallyspProAlaPheGluGluPheCysLeuProThrIysTyrAlaAspVal 208
Db 848 ACGTTCGTCAGCCTGCTTGGAGAAATCTGCTTGCCACAAAGAAAGTATGCTGATGTG 907

Qy 209 IleIleProArgGly 213
Db 908 ATCATCCCTAGAGGT 922

RESULT 12
US-09-908-975-398
; Sequence 398, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 398
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-908-975-398

Alignment Scores:
Pred. No.: 2.68e-07 Length: 65
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-09-908-975-398 (1-65)

Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn 232
Db 20 GTGGCATTCACTCATCTGTCGACGATATACGGACATCTCTAATC 64

RESULT 13
US-09-908-975-24159
; Sequence 24159, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
```

APPLICANT: FAIGLER, Simchon
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
FILE REFERENCE: 36686-0005
CURRENT APPLICATION NUMBER: US/09/908,975
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24159
LENGTH: 65
TYPE: DNA
ORGANISM: Mus musculus
US-09-908-975-24159

Alignment Scores:
Pred. No.: 2,686-07 Length: 65
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x US-09-908-975-24159 (1-65)

OY 160 ThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
Db 1 ACAGACTCTGATGTAGGCTGTCTCGAAGAGTTCTCCGGATGTG 45

RESULT 14
US-10-066-543-1843
Sequence 1843, Application US/10066543
Publication No. US2003087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yuhui
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indritas, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Derrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1843
LENGTH: 447
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1843

Alignment Scores:
Pred. No.: 1,526-06 Length: 447
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 14 Gaps: 0

US-09-896-522-2 (1-277) x US-10-066-543-1843 (1-447)

OY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyIysSer 37
|||||

Db 71 CCTTCTATTAGCGCTCAGCGGGGACACGACTAGCGCAAGTCT 115

RESULT 15
US-09-918-995-23923
Sequence 23923, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23923
LENGTH: 455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(455)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

Alignment Scores:
Pred. No.: 1,546-06 Length: 455
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-23923 (1-455)

OY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyIysSer 37
Db 85 CCTTCTATTAGCGCTCAGCGGGGACACGACTAGCGCAAGTCT 129

Search completed: November 25, 2003, 12:17:21
Job time : 379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:12:16 ; Search time 362 Seconds
(without alignments)
2521.075 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450

Sequence: 1 MASAGEDCEPAPRAPH.....HPGMLTSGKSHLESSRPH 277

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2190069 seqs, 1647345023 residues
Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09896522/runtac.21112003.184356.5076/app_query.fasta.1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.csl -LIST=45 -DOCLALIGN=200 -THR SCORE=oct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09896522 @cgn_1.1.347 @runtac.21112003.184356.5076
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PCFUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq: *
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	834	9	US-09-896-522-3 Sequence 3, Appl1

2	1450	100.0	1288	14	US-10-037-270-546	Sequence 546, App
3	1450	100.0	1624	9	US-09-896-522-1	Sequence 1, Appl1
4	1437	99.1	1648	10	US-09-833-381-2048	Sequence 2048, Ap
5	940	64.8	1402	13	US-10-098-841-53	Sequence 53, Appl
6	902.5	62.2	1310	10	US-09-925-300-220	Sequence 220, App
7	727	50.1	472	11	US-09-918-995-30379	Sequence 30379, A
8	647	44.6	19125	10	US-09-764-877-2805	Sequence 2805, Ap
9	544	37.5	1401	10	US-09-938-842A-1943	Sequence 1943, Ap
10	540.5	37.3	2058	12	US-09-918-995-24042	Sequence 24042, A
11	537.5	37.1	472	11	US-10-066-543-1843	Sequence 1843, Ap
12	521.5	36.0	447	14	US-09-918-995-23923	Sequence 23923, A
13	511.5	35.3	455	11	US-09-070-927A-215	Sequence 215, App
14	427	29.4	3662	10	US-10-329-960-1	Sequence 960, App
15	384	26.5	1830121	14	US-09-918-995-9667	Sequence 9667, App
16	372.5	25.7	470	11	US-09-960-885-11	Sequence 11, Appl
17	351.5	24.2	15051	12	US-09-960-870-11	Sequence 11, Appl
18	351.5	24.2	15051	12	US-10-205-220-1	Sequence 1, Appl1
19	351.5	24.2	580073	12	US-10-029-366-15883	Sequence 15883, A
20	327	22.6	187	12	US-10-029-366-2183	Sequence 2183, Ap
21	327	22.6	510	12	US-09-764-877-2804	Sequence 2804, Ap
22	327	22.6	9732	10	US-08-781-986A-955	Sequence 955, App
23	212	14.6	514	8	US-09-938-842A-515	Sequence 615, App
24	205.5	14.2	1188	10	US-09-770-445-403	Sequence 403, App
25	199.5	13.8	926	9	US-09-974-300-626	Sequence 626, App
26	194.5	13.4	253	10	US-09-974-300-609	Sequence 609, App
27	177.5	12.2	942	10	US-09-923-876-5130	Sequence 5130, Ap
28	175.5	12.1	328	10	US-09-923-876-5906	Sequence 5906, Ap
29	174.5	12.0	263	9	US-09-738-626-1109	Sequence 1109, Ap
30	173	11.9	248	9	US-09-738-626-1	Sequence 1, Appl1
31	154.5	10.7	3309400	10	US-08-781-986A-995	Sequence 995, App
32	154.5	10.4	9025608	14	US-10-156-761-1	Sequence 1, Appl1
33	150.5	10.3	729	8	US-10-329-960-1	Sequence 298, App
34	149.5	10.3	967	14	US-09-878-574-4	Sequence 4, Appl1
35	149	10.3	987	14	US-09-734-569-87	Sequence 87, Appl
36	137.5	9.5	1830121	14	US-09-736-457-1476	Sequence 1476, Ap
37	136.5	8.5	4029	10	US-09-902-941-1476	Sequence 1476, Ap
38	123.5	8.4	339	10	US-09-849-626-1476	Sequence 1476, Ap
39	119.5	8.2	433	9	US-10-113-872-1476	Sequence 1476, Ap
40	117	8.1	401	10	US-10-017-754-1476	Sequence 1476, Ap
41	117	8.1	401	10	US-09-938-842A-2093	Sequence 2093, Ap
42	117	8.1	401	12		
43	117	8.1	401	12		
44	117	8.1	401	14		
45	111.5	7.7	1425	10		

ALIGNMENTS

RESULT 1
US-09-896-522-3
; Sequence 3, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896, 522
; PRIOR APPLICATION NUMBER: 2001-06-28
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-896-522-3

Alignment Scores:
Pred. No.: 2,81e-175
Score: 1450.00
Percent Similarity: 100.00%

Length: 834
Matches: 277
Conservative: 0

QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgLeu 180
 Db 575 GACTCCGACGTGAGGCTGTCTCGAAGATTCTCCGGACGTGCGCGAGGAGGACTG 634
 QY 181 GluGlnIleuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheVal 200
 Db 635 GAGCAGATTCTGACGCGACGTACACACCTTCGTGAGAGCGGCTTCGAGGATTCTGCTG 694
 QY 201 ProThrIysIysTyrAlaAspValIleIleProArgIlyValAspAsnMetValAlaIle 220
 Db 695 CCGACAAAGAGATGCGCGATGATCATCCACGAGAGTGGACAAATATGTTGCCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTrpHisArg 240
 Db 755 AACCTGATCGTGACGACATCCAGGACATTCGTGATGTGACATCTGCAATGGACCGA 814
 QY 241 GlyGlySerAsnGlyArgSerTyrIlyAsArgThrPheSerGluProGlyAspHisProGly 260
 Db 815 GAGGGGTCCATGCGCGACGTACCAAGCGGACCTTTCTAGCCAGGAGGACCACTGGG 874
 QY 261 MetLeuThrSerGlyIlyAsArgSerHisLeuGluSerSerArgProHis 277
 Db 875 ATGCTGACCTCTGGCAAGCGTCACTTTGAGTCCAGCAGAGACCCAC 925

RESULT 3

US-09-896-522-1
 ; Sequence 1, Application US/09896522
 ; Patent No. US20020055161A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria A.
 ; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 381552001700
 ; CURRENT APPLICATION NUMBER: US/09/896,522
 ; CURRENT FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: 60/216,503
 ; PRIOR FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1624
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (94)...(927)
 US-09-896-522-1

Alignment Scores:

Pred. No.: 7,76e-175 Length: 1624
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x US-09-896-522-1 (1-1624)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 94 ATGGCTTCGGCGGAGCGGAGCGAAGACCTCGAGAGCCCGCGCGGAGCGCGCTCCGAC 153
 QY 21 GlnArgProPheLeuIleGlyValAsnGlyTyrThrAlaSerGlyIysSerThrValCys 40
 Db 154 CAGCGGCGCTTCCTGTAAGGGGTGAGCGGCGGCACTGCGCGGGAAGTGCAGCCGTGCT 213
 QY 41 GluIlyIleMetCysLeuLeuGlyGlnAsnGlyValGluGlnArgGlnArgIysValVal 60
 Db 214 GAGAGATCATGAGTTCCTGCGGACAGAAAGAGTGGAAAGCGCGACCGGAGAGTGTGTC 273
 QY 61 IleLeuSerGlnAspArgPheTyrIlyAsArgThrAlaGluGlnIlyAsnIlyAsnAlaLeu 80
 Db 274 ATCTTGAGCCAGACAGGTTCTCAAGGTCTGACGCGAGAGCAAGGCAAGGCGCTTG 333

QY 81 IysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 334 AAGAGCAGTACATTTTGAACATCCAGATGCTTTGATATATATTTGATGACAGAGACT 393
 QY 101 LeuIlyAsnIleValGluGlyIlySerValGluValProThrTyrAspPheValIleHis 120
 Db 394 CTGAGAGACATCGTGGAGGCAAAAGCGTGGAGGTCCGACCTATATATTTTGGACACAC 453
 QY 121 SerArgLeuProGluThrThrValIlyTyrProAlaAspValIleuPheGluGlyIle 140
 Db 454 TCAAGTTACAGAGACCAAGCGGTGTCTACCTCCGCGAGCGTGTCTGTTTGGAGGCATC 513
 QY 141 LeuValPheTyrSerGlnGlnIlyLeuAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 514 TTGGTGTTCACAGCCAGGAGATCCGGACATGTTCCACCTGCGCTTCCTGTGGACACC 573
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgIlyArgAspLeu 180
 Db 574 GACTCCGACGTGAGGCTGTCTCGAAGATTCTCCGGACGTGCGCGAGGAGGACTG 633
 QY 181 GluGlnIleuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
 Db 634 GAGCAGATTCTGACGCGACGTACACACCTTCGTGAGAGCGGCTTCGAGGAGTCTGCTG 693
 QY 201 ProThrIlyIysTyrAlaAspValIleIleProArgIlyValAspAsnMetValAlaIle 220
 Db 694 CCGACAAAGAGATGCGCGATGATCATCCACGAGAGTGGACAAATATGTTGCCATC 753
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTrpHisArg 240
 Db 754 AACCTGATCGTGACGACATCCAGGACATTCGTGATGTGACATCTGCAATGGACCGA 813
 QY 241 GlyGlySerAsnGlyArgSerTyrIlyAsArgThrPheSerGluProGlyAspHisProGly 260
 Db 814 GAGGGGTCCATGCGCGACGTACCAAGCGGACCTTTCTGAGCAGGAGGACCACTGGG 873
 QY 261 MetLeuThrSerGlyIlyAsArgSerHisLeuGluSerSerArgProHis 277
 Db 874 ATGCTGACCTCTGGCAAGCGTCACTTTGAGTCCAGCAGAGACCCAC 924

RESULT 4

US-09-833-381-2048
 ; Sequence 2048, Application US/09833381
 ; Patent No. US20020132090A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: NO. US20020132090A1 Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; CURRENT FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2048
 ; LENGTH: 1648
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(1648)
 ; OTHER INFORMATION: n = A, T, C or G
 US-09-833-381-2048

Alignment Scores:

Pred. No.: 3,64e-173 Length: 1648
 Score: 1437.00 Matches: 277
 Percent Similarity: 99.64% Conservative: 0
 Best Local Similarity: 99.64% Mismatches: 1
 Query Match: 99.10% Indels: 0
 DB: 10 Gaps: 0


```

Qy 163 AspValArpheuserArgValIleuArgAspVal---ArgArgIYARgAspLeuGlu 101
Db 767 GACACCCGGCTCTCACGACAGATTAAGGACATCCGACAGAGAGCGAGATCTTGAG 826
Qy 182 GlnIleLeuThrgInTyThrThrPheValIysProAlaPheGluGluPheCysLeuPro 201
Db 827 CAGATTTTACTGATGATTAAGTCTTGTCAAGCTCTCTTGAAGAACTCTGCTTCCA 866
Qy 202 ThrIysLysTyraIaAspValIleIleProArgGlyValaIaphenMetValaIaIaAsn 221
Db 887 ACAAGAAGATGATGTGATGATCATCTTGAAGTGCAGATATATCTGTGGCCATCAAC 946
Qy 222 LeuIleValGlnHisIleGlnaPheIleLeuAsnGlyAspIleCysLysTrpHisArgGly 241
Db 947 CTCTATCGTCAGACATCCAGACATCTTAAGTGA----- 982
Qy 242 GlySerAsnGlyArgSerTyraIysArgThrPheSerGluProGlyAspHisProGlyMet 261
Db 983 -----GGGCCCTCCAAACGCGACACC-----AATGGCTGTCTCAACGCGCTAC 1024
Qy 262 LeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
Db 1025 ACCCCTTCACGACAGAGCGAGCATCGAGTCCAGACGACGCGCGCAT 1072

RESULT 6
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-220

Alignment Scores:
Pred. No.: 5,52e-105 Length: 1310
Score: 902.50 Matches: 189
Percent Similarity: 77.82% Conservative: 32
Best Local Similarity: 66.55% Mismatches: 40
Query Match: 62.24% Indels: 24
Gaps: 5

US-09-896-522-2 (1-277) x US-09-925-300-220 (1-1310)
Qy 3 SerAlaGlyGluLysProCysGluSerProAlaProGluAlaAspArgPro----- 19
Db 205 GCAGCGGGAGAGAGGGCGGCGGACATGGCCGGGACGACGACGACGACCTTCGACGAC 264
Qy 20 HisGlnArg-----ProPheLeuIleGlyValaSerGlyGlyThraIaSer 34
Db 265 CACCACACGACCAACGCGCGGCGGACCTTCTTATAGCGCTCAC-GGGGGAACAGCTAGC 323
Qy 35 GlyLysSerThraIaCysGluLysIleMetGluLeuLeuGluGlnaGlnIaIaGlu 54
Db 324 GGCAGAGCTTCCTGCTGTGCTAAGATGTCAGAGCTCTGGGGGAGATGAGGTGACTAT 383
Qy 55 ArgGlnArgLysValaIleLeuSerGlnaAspArgPheTyraIysValaLeuThraIaGlu 74
Db 384 CGCCAGAGAGAGGTGTCTGATCTGAGCCAGATAGCTTTCACGCTGTCTTACCTCGAG 443

```

```

Qy 75 GlnIysAlaIysAlaLeuLysGlyGlnTyraAsnPheAspHisProAspAlaPheAspAsn 94
Db 444 CAGAGGCGCAAGAGCCCTGAAGGSCCATTTCAATTGACACCCGGATGCTTTGACAAAT 503
Qy 95 AspLeuMetHisArgThrLeuLysAsnIleValaGluGlyLysThraIaGluValaProThr 114
Db 504 GASTGATCTCAAAACATCAAAAGAAATCTACGAAAGGAGAAACATCCAGATCCCGGTG 563
Qy 115 TyrAspPheValThrHisSerArgLeuProGluThrThraValaIyrrProAlaAspVal 134
Db 564 TATGACTTGTCTCCCATTCGCCGAAGAGAGACATTACTGTCTATCTCCGACGACGTG 623
Qy 135 ValLeuPheGluGlyIleLeuValaPheTyraSerGlnGluIleArgAspMetPheHisLeu 154
Db 624 GTCCTTTTGAAGGAGATCCCTGGCTTCTACTCCAGAGAGATACGACCTGTTCAGATG 683
Qy 155 ArgLeuPheValaAspThrAspSerAspValArgLeuSerArgArgValaLeuArgAspVal 174
Db 684 AACCTTTTGTGATACAGATGCGGACACCCGCTCTCACGACGATATTAAGGACATC 743
Qy 175 --ArgArgGlyArgAspLeuGluGlnIleLeuThrgInTyThrThrPheValLysPro 193
Db 744 AGCGAGAGAGCGAGGATCTTGACAGATTTTATCTCAGTACATTCAGTTCCGCAAGCCT 803
Qy 194 AlaPheGluGluPheCysLeuProThraIysLysTyraIaAspValIleIleProArgGly 213
Db 804 GCCTTTGAGAAATTCGCTTGCCCAAGAGATGATGTGATGTATCTCTTAAGGT 863
Qy 214 ValaAspAspMetValaIleLeuLeuIleValaGlnHisIleGlnaPheIleLeuAsnGly 233
Db 864 GCAGATTAATCTGTGGCCATCAACCTCATCTGTCGACACATCCAGACATCTCGAATGCA 923
Qy 234 AspIleCysLysTrpHisArgGlyLysSerAsnGlyArgSerTyraIysArgThrPheSer 253
Db 924 -----GGGCCCTCCAAACGCGACACC----- 944
Qy 254 GluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSer 273
Db 945 ---AATGGCTGTCTCAACGGCTTACACCCCTTCACGCAAGAGAGGAGCATCGAGTCCAGC 1001
Qy 274 SerArgProHis 277
Db 1002 AGCAGGCGCGCAT 1013

RESULT 7
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995.
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)---(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Alignment Scores:
Pred. No.: 3.18e-83 Length: 472
Score: 727.00 Matches: 143

```

Percent Similarity: 99.31% Conservative: 0
 Best Local Similarity: 99.31% Mismatches: 1
 Query Match: 50.14% Indels: 0
 DB: 11 Gaps: 0

US-09-896-522-2 (1-277) x US-09-918-995-30379 (1-472)

QY 33 AAserglylyserThValCysglylylleuMetguleuLeuGlylnAsnGluVal 52
 Db 41 GCCAGCGGAGGAGTCCAGCTGTGTGAGAGATCATGAGTTGCTGGACAGAGAGGTG 100
 QY 53 GluGlnArgGlnArgValValIleleuSerGlnAspArgPheTyrValIleuThr 72
 Db 101 GAACAGCGGACGCGAAGGTGATCTCTGAGCCGAGCAGAGTCTTACAAAGTCTGACG 160
 QY 73 AlaGluGlnlyslAlaValAlaLeuLysGlyIleTyrAsnPheAspHisProAspAlaPhe 92
 Db 161 GCAAGCGAAGGCGCAAGGCTTGAAGACAGTACATTTGACCATCCAGATGCCCTTT 220
 QY 93 AspaAsnAspLeuMetHisArgThrIleuLysAsnIleValGluGlyThrValGluVal 112
 Db 221 GATATGATTTGATGACAGACATCTGAGAACATCGTGGAGGCAAAACGTTGAGGTG 280
 QY 113 ProThrTyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAla 132
 Db 281 CCGACCTATGATTTGTGTGACACACTCAAGGTTACAGAGACACAGGTGTCTACCCGTGCG 340
 QY 133 AspValValIleuPheGluGlyIleleuValPheTyrSerGlnGluIleArgAspMetPhe 152
 Db 341 GACGTGTTCTGTTGTGAGGCGATCTGTGTTCTTACAGCCAGAGATCCGGGACATGTTCC 400
 QY 153 HisleuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValIleuArg 172
 Db 401 CACCTGCGCCTCTTCGTGTGACACCGACTCCACGCTCAGGCTGTATCGAAGGTTCTCCGG 460
 QY 173 AspValArgArg 176
 Db 461 GACGTGCGCCGA 472

RESULT 8

US-09-764-877-2805
 ; Sequence 2805, Application US/09764877
 ; Patent NO. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764,877
 ; PRIORITY FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2805
 ; LENGTH: 19125
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-877-2805

Alignment Scores:

Pred. No.: 1,53e-70 Length: 19125
 Score: 647.00 Matches: 197
 Percent Similarity: 26.51% Conservative: 1
 Best Local Similarity: 26.37% Mismatches: 3
 Query Match: 44.62% Indels: 548
 DB: 10 Gaps: 4

US-09-896-522-2 (1-277) x US-09-764-877-2805 (1-19125)

QY 1 MetAlaSerAlaGlyGluAspCysGlnSerProAlaProGluAlaAspArgProHis 20
 Db 41 ATGGCTTCGGCGGAGGCGAAGACTCGAGAGCCCGCGCGGCGGAGCCGTCGCCAC 100
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLys----- 36

Db 101 CAGGGGCGCTCTTGATAGGGGTGAGGGCGGACATGCGAGCGGGAAGTAAGGGGCCGA 160
 QY 36 ----- 36
 Db 161 GCGGGCGGGGTCTGCTCTCTCGCGGTGCTTACGGGCGGCGCGGTCCAGAGAGAGACC 220
 QY 36 ----- 36
 Db 221 CAGACCCCGCCAGAGGGCCCCCTTCCAGGACAGGGGCGAGGCGGAGGAGGGCTG 280
 QY 36 ----- 36
 Db 281 CAGGCGCATGAGGCGGGCGGGCTGGGGCGGCGAGGGGTTGGGGCGCTGGGGCTCGG 340
 QY 36 ----- 36
 Db 341 GGCTCCCGCCACAGCCTTCTCCGCGGGCGGCTCTCGAGCCCAAGTCCGTCTAACCC 400
 QY 36 ----- 36
 Db 401 CGTTGGGCGCGCGGCGCTTCCGTGAGACATCCACACTGTGTGTTCTCCGCGCT 460
 QY 36 ----- 36
 Db 461 TCGAGCATACCCCCAGGGGCGGGTGGGTCAGTGTTCCAGAACCGGCGCGTGCCT 520
 QY 37 -----SerThrValCy 40
 Db 521 GTCAAGAGACAGCGGTTCCAGCGGGGAATCCCGCTGTGTCTTCAGTCGACCTGTG 580
 QY 40 sglLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
 Db 581 TGACAAATATGATGAGTTGCTGTGGAACAGAGGTGAGAACAGCGGACAGCGAGGTGT 640
 QY 60 lIleuSerGlnAspArgPheTyrLysValleuThrAlaGluGlnLysAlaLysAla 80
 Db 641 CATCTGAGCCAGACAGCGGTTCTTACAAAGTCTGAGCGGACAGACAGGCGCAAGCGCTT 700
 QY 80 uLysGlyGlnTyrAsnPheAspHisPro----- 88
 Db 701 GAAAGGACAGTACAAATTTTGAACATCCAGTAATCGAGGCCAAGCCATGTGGAGAGG 760
 QY 89 ----- 89
 Db 761 CAGACCCCATTTGCGGCGCGGCTTAACAGCTGCGCTTGCGGCCCAAGTTGATGATGC 820
 QY 89 ----- 89
 Db 821 AGACCGGCTTGGGGTGTTCCTCCCAACTCGCCACAGAGCTGCAGTGAGTGTGCTGCA 880
 QY 89 ----- 89
 Db 881 CAGTGAATCAGGGAACCTTGTGAGTTCCTGAGCCCTGAGTGTCCCCCAGAGCGT 940
 QY 89 ----- 89
 Db 941 GGTGAGCAGTCAGAGAAAGCTGTTCACATGTGAGGAGAGCCCTGCTGCTGCTGCA 1000
 QY 89 ----- 89
 Db 1001 CTGTGATCTGAGAAATCTCCCATCTTCCCTTCCAGAGTGGGGCTTTAGGGCGCTC 1060
 QY 89 ----- 89
 Db 1061 CACACACACTCTGCCCACTGTGCTCCCGTCTTTGGGGCAAAAGTCTCAGAGAGGCC 1120
 QY 89 ----- 89
 Db 1121 AGTGTGAGAGGCGGTGTGTAACTAAGTGTGTGCTTGAAGTCCGGGCGAGGTAAAG 1180
 QY 89 ----- 89

Db 1181 GGCAGATGCGGACGAGAGACACCTGACCTAGTAGAAGATTCCACCTGGCTGGCTGGCA 1240
Qy 89 ----- 89
Db 1241 TGGTCAGCCCTTTTCCCAAGGACTTGTGGAGAAAGCGCTGCAGAGCCGAGCTTAACGAG 1300
Qy 89 ----- 89
Db 1301 TTAGTCCCAAGAAAGAGAGAGAGAGAGGCGCCAGCCTTACCACCTCCAGAAAGAGCTGA 1360
Qy 89 ----- 89
Db 1361 TCAGAAAACCCCTGTGTTCGCGCATCTCAGGGGACAGAGCTTGGGCGAGTGACAGACCC 1420
Qy 89 ----- 89
Db 1421 GAGAAACAG 1480
Qy 89 ----- 89
Db 1481 CCAAACTACACTGCGCTAGTCTAGTCTAGTCTAGGCTTCTGCTGCTGCTGCTGCTGCTG 1540
Qy 89 ----- 89
Db 1541 CCCAGCTGTGCTGACCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1600
Qy 90 ----- 90
Db 1601 TGCCTGTCTTGCCTTGTGTCTCCAACTCTGAGATGCTTGTGATTAATGATTTGATGCA 1660
Qy 99 GTTtLeuLyAsn11eValGluGlyLysThr-ValGluValProThrTyrAspPheValT 119
Db 1661 GACTCTGAACACATCTGTGAAGGAGCAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1720
Qy 119 hrHisSer----- 121
Db 1721 CACACTC-AAGGTAAAG 1779
Qy 121 ----- 121
Db 1780 GTCCCTCGGGCTCACACAGACCTGCGCGAGAGCCCTCCCGGTGGTTCTGTGAG 1839
Qy 121 ----- 121
Db 1840 GTCCCCACCCCTTCCACAGTACGTCTCCCTGCATGAACTGACTGTGTGTGGCC 1899
Qy 122 ----- 126
Db 1900 TGGAGACCTGTGTCTGACAGAGAGCCCTTCCCGGTGGCCCTCAGGTTTACAGAGA 1959
Qy 126 hrThrValValTyrProAlaAspValValLeuPheGluGly11eLeuValPheTyrSerG 146
Db 1960 CCAAGGTGTCTACCTCCGAG 2019
Qy 146 lnglu1leaArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgL 166
Db 2020 AGAGAGATCCGGAGCATGTTCACCTGCGCTTTCGTGACACCGACTCCGAGGTGAGGC 2079
Qy 166 euSerArgArgVal----- 170
Db 2080 TGTCTCGAAGAGG-TAAGGCGCGCGCGGCGCTCCGTGCTGCGCCGAGGCTGGCGGC 2138
Qy 171 ----- 176
Db 2139 AGGTCTGAGGTGTGAAGACAGCGCTGCGCTTGGCCGTGAGGTCTCCCGGAGAGGTGGCC 2198
Qy 176 rgglYArgAspLeuGluGln11eLeuThrGlnTyrThrThrPheValValProAlaPheG 196
Db 2199 GAGGAG 2258
Qy 196 lngluPheCyLeuPro 201
Db 2259 AGGAGTTCTGCTGCGG 2275

RESULT 9
US-09-938-842A-1943
; Sequence 1943, Application US/09938842A
; Patent No. US20020160378B1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1943
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1943
Alignment Scores:
Pred. No.: 4.16e-59 Length: 1401
Score: 544.00 Matches: 116
Percent Similarity: 63.16 Conservative: 52
Best Local Similarity: 43.618 Mismatches: 78
Query Match: 37.528 Indels: 20
Gaps: 8
US-09-896-522-2 (1-277) x US-09-938-842A-1943 (1-1401)
Qy 11 SerProAlaProGlu-----AlaSerArgProHisGlnArgProPheLeu11egly 27
Db 85 TCMAATAGACACAGAGAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 144
Qy 28 ValSerGlyGlyThrValSerGlyLysSerThrValCysGluLys11eMetGluLeu 47
Db 145 GTTGTGAGGTGACACATCCGAGAAACACTGTGTGATGATGATGATGATGATGATGATGATGAT 204
Qy 48 GlyGlnAsnGluValGluGlnArgGlnArgLysValVal11eLeuSerGlnAspArgPhe 67
Db 205 -----CATGATCAGAGAGCGTGTGTGTAATCAGGATTCCTTC 243
Qy 68 TyrLeuValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAsp 87
Db 244 TACCATTAATGTAAATGAACTGAGAGCTTGTGAAGATT-----CATGATTAACAATTTTGAC 297
Qy 88 HisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsn11eValGluGly 107
Db 298 CATCTGACGCTTTCATCTGACATCTGACATTAATTTCTCTTCATGAGAGAGTTAAGAAAGGA 357
Qy 108 LysThrValGluValProThrTyrAspPheValThrHisSerArg-----LeuProGlu 125
Db 358 CAAGCAGTGAATATTCCTTAATTAATGACTTCAAAAGTTACAAAGAACAAATGTTTTCACCT 417
Qy 126 ThrThrValValTyrProAlaAspValValLeuPheGluGly11eLeuValPheTyrSer 145
Db 418 AGAAGG---GTAAATCTCTTGAAGAGTTAATTTGGAAGGTAAATCTCAATTTTCCATGAC 474
Qy 146 GlnGlu1leaArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArg 165
Db 475 CCTCGGAGTCCGAGATTTGATGAACATGAAGATATTTGTGAAGAGAGAGAGAGAGAGAGAG 534
Qy 166 LeuSerArgArgValLeuArgAsp---ValArgArgGlyArgAspLeuGln11eLeu 184
Db 535 CTAGCGAAGAAAGATTAAGCGCATACTGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594

Qy 185 ThrGlnTyrThrPheValIlyserProAlaPheGlnGluPheCysIleuProThrIlySer 204
 Db 595 GACCACTCTCAAGATTGTGTGAAGCCACATTGAGATTTCATACCTCCACCAAGAA 654
 Qy 205 TyrAlaaspValIleIleProArgIlyValAspAsnMetValAlaIleAsnIleVal 224
 Db 655 TACCAAGAT 714
 Qy 225 GlnHisIleGlnAspIleLeu---AsnGlyAspIleCysIlySerIlyPheArgIlySer 243
 Db 715 CAACACATCCATACAGAGCTGGTCAACATATCTGTAAATTAT---CCA 765
 Qy 244 AsnGlyArgSerIlyArgThrPheSerGluProGlyAspHisProGlyMetIleThr 263
 Db 766 AATCTTAT 825
 Qy 264 SerGlyArgSerHis 269
 Db 826 TCTAAACACAAAGCAT 843

RESULT 10

US-10-251-186-14
 ; Sequence 14, Application US/10251186
 ; Publication No. US20030180745A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030180745A1 Nucleic Acids and
 ; FILE REFERENCE: 789CIP2CDV1
 ; CURRENT APPLICATION NUMBER: US/10/251,186
 ; PRIOR FILING DATE: 2002-09-19
 ; PRIOR APPLICATION NUMBER: 09/665,363
 ; PRIOR FILING DATE: 2000-09-19
 ; PRIOR APPLICATION NUMBER: 09/574,454
 ; PRIOR FILING DATE: 2000-05-19
 ; PRIOR APPLICATION NUMBER: 09/519,705
 ; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: pc_fl_genes Version 2.0
 ; SEQ ID NO 14
 ; LENGTH: 2058
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (173)..(1909)
 US-10-251-186-14

Alignment Scores:

Pred. No.: 2,1e-58 Length: 2058
 Score: 540.50 Matches: 110
 Percent Similarity: 64.20% Conservative: 46
 Best Local Similarity: 45.27% Mismatches: 60
 Query Match: 37.28% Indels: 27
 DB: 12 Gaps: 5

US-09-896-522-2 (1-277) x US-10-251-186-14 (1-2058)

Qy 9 CysGluSerProAlaPro-----GluAlaAspArg 18
 Db 455 TGCAGTCAAGAGCTCCCTCTGCTACAGCAAGCTACATCAACCGCGCGG 514
 Qy 19 Pro-----HisGlnArgProPheLeuIleGlyValSer 29
 Db 515 CCGGCTGTGTAATGAACAGCAGCAATCAAGAGGCTTGGCATCGGCTTGGGA 574
 Qy 30 GlyGlyThrAlaSerGlyIlySerThrValCysGluIlyIleMetGluLeuGln 49
 Db 575 GCGCGAGTGTCTCTGGGAAACCACTGTGCGCAGAAATGATCATGAGCGCTG----- 628

Qy 50 AsnGluValGlnGlnArgGlnArgIlyValIleIleSerGlnAspArgPheIlySer 69
 Db 629 ---GATGTGCTCGG-----GTGGTCTTGTCTGCTCATGTGCTCTTACAG 673
 Qy 70 ValIleuThrAlaGlnGlnIlyValAlaIleIlyGlyGlnIlyPheAsnPheAspHisPro 89
 Db 674 GTGCTGACTGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 723
 Qy 90 AspAlaPheAspAsnIleuMetHisArgThrIleIlyAsnIleValGlnGlyIlyThr 109
 Db 734 GATGCTTGTACTTCACTCATCATCTTCCACCTCAAGAGGCTGAAGGAGGAGGAGGAG 793
 Qy 110 ValGluValProThrIlyAspPheValThrHisSerArgIlyProGluThrThrVal 129
 Db 794 GTCAAGTGTCCCTTTATATGCTTCCACAGCCAGCCGGAAGAAGACTGGAACACTG 853
 Qy 130 TyrProAlaAspValValIleuPheGlnGlyIleLeuValPheTyrSerGlnIlyIleArg 149
 Db 854 TATGCTGCAAAACGTCATCATCTTGAAGGAGCATGAGCCTTGTCTGACAAAGACTGTTG 913
 Qy 150 AspMetPheHisIleuArgIlyPheValAspThrAspSerAspValArgIlySerArgArg 169
 Db 914 GAGCTCTGAGCATGAAGATCTTGTGACACAGACTCCGACATCCGCTGTACGCGG 973
 Qy 170 ValIleuArgAspVal---ArgArgGlyArgAspIlyGlnIleIleuThrGlnTyrThr 188
 Db 974 CTGCGCCGAGCATCACTGAGCCGCGCGGAGCATGAGGCTGTCAACAGCTACAC 1033
 Qy 189 ThrPheValIlyProAlaPheGlnGluPheCysIleuProThrIlyIlySerIlyAlaAspVal 208
 Db 1034 AAGTTGTCAAGCCTCTTGCAGCATCATCAGCCACCATGCGCGCTGGCAGACATC 1093
 Qy 209 IleIleProArgIlyValAspAsnMetValAlaIleAsnIleIleValGlnHisIleGln 228
 Db 1094 GTGCTCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1153
 Qy 229 AspIleLeu 231
 Db 1154 AGCCAGCTG 1162

RESULT 11

US-09-918-995-24042
 ; Sequence 24042, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 09/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 24042
 ; LENGTH: 472
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(472)
 ; OTHER INFORMATION: n = A, T, C or G
 US-09-918-995-24042

Alignment Scores:

Pred. No.: 5.36e-59 Length: 472
 Score: 537.50 Matches: 103
 Percent Similarity: 84.14% Conservative: 19
 Best Local Similarity: 71.03% Mismatches: 22
 Query Match: 37.07% Indels: 1
 DB: 11 Gaps: 1

US-09-896-522-2 (1-277) x US-09-918-995-24042 (1-472)

```
Qy      8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
      |||
      37 GACAGCGAGACGACCCCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATATA 96
Qy      27 G1yValSerG1yG1yThraAlaSerG1yLysSerThraValCysG1uLysIleMetG1uLeu 46
      97 GGGCTACGGCGGGGGAACAGCTACGCGCAAGCTTCCGTGTGTCTTAAGATCGTCAGCTC 156
Qy      47 LeuG1yGlnAsnG1uValG1uGlnArgGlnArgLysValValIleLeuSerGlnAspArg 66
      157 CTGGGGCAGATAGGTGAGCTATCGCCAGACGAGTGTGCATCTGAGCCAGAGATAGC 216
Qy      67 PheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysG1yGlnTyrAspPhe 86
      |||
      217 TTCTACCGGTCTCTTACTCGGAGCAAGAGCCCAAGCCCTGAAAGGCGCAGTTCACTTT 276
Qy      87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGlu 106
      |||
      277 GACCACCCGATGCTTGTGACATGAACTATTCTCAAAACACTCAAAAGAAATCACTGAA 336
Qy      107 G1yLysThraValGluValProThrTyrAspPheValThriHisSerArgLeuProGluThr 126
      337 GGGAAAACAGTCCAGATCCCGGTATGACTTGTCTCCCATTCGCGAAGAGAGAGACA 396
Qy      127 ThrValValTyrProAlaAspValValLeuPheGluG1yIleLeuValPheTyrSerGln 146
      397 GTTACTGTCTATCCCGCAGAGAGTGCTCTTTGAAAGGAGATCCGCTTACTCCAG 456
Qy      147 G1uIleArgAspMet 151
      |||
      457 GAGGTACGAGACCTG 471
Db
```

RESULT 12
US-10-066-543-1843
Sequence 1843, Application US/10066543
Publication No. US20030087818A1

GENERAL INFORMATION:
APPLICANT: Jiang, Yuguo
APPLICANT: Pye, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1843
LENGTH: 447
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1843

Alignment Scores:

Pred. No.:	5,47e-57	Length:	447
Score:	521.50	Matches:	101
Percent Similarity:	83.69%	Conservative:	17
Best Local Similarity:	71.63%	Mismatches:	22
Query Match:	35.97%	Indels:	1
DB:	14	Gaps:	1

US-09-896-522-2 (1-277) x US-10-066-543-1843 (1-447)

```
Qy      8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
      |||
      23 GACAGCGAGACGACCCCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATATA 82
Qy      27 G1yValSerG1yG1yThraAlaSerG1yLysSerThraValCysG1uLysIleMetG1uLeu 46
      83 GGGCTACGGCGGGGGAACAGCTACGCGCAAGCTTCCGTGTGTCTTAAGATCGTCAGCTC 142
Qy      47 LeuG1yGlnAsnG1uValG1uGlnArgGlnArgLysValValIleLeuSerGlnAspArg 66
      143 CTGGGGCAGATAGGTGAGCTATCGCCAGACGAGTGTGCATCTGAGCCAGAGATAGC 202
Qy      67 PheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysG1yGlnTyrAspPhe 86
      |||
      203 TTCTACCGGTCTCTTACTCGGAGCAAGAGCCCAAGCCCTGAAAGGCGCAGTTCACTTT 262
Qy      87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGlu 106
      263 GACCACCCGATGCTTGTGACATGAACTATTCTCAAAACACTCAAAAGAAATCACTGAA 322
Qy      107 G1yLysThraValGluValProThrTyrAspPheValThriHisSerArgLeuProGluThr 126
      323 GGGAAAACAGTCCAGATCCCGGTATGACTTGTCTCCCATTCGCGAAGAGAGAGACA 382
Qy      127 ThrValValTyrProAlaAspValValLeuPheGluG1yIleLeuValPheTyrSerGln 146
      383 GTTACTGTCTATCCCGCAGAGAGTGCTCTTTGAAAGGAGATCCGCTTACTCCAG 442
Qy      147 G1u 147
      |||
      443 GAG 445
Db
```

RESULT 13
US-09-918-995-23923
Sequence 23923, Application US/09918995
Publication No. US20030073623A1

GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23923
LENGTH: 455
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(455)
OTHER INFORMATION: n = A, T, C or G
US-09-918-995-23923

Alignment Scores:

Pred. No.:	1.07e-55	Length:	455
Score:	511.50	Matches:	99
Percent Similarity:	83.45%	Conservative:	17
Best Local Similarity:	71.22%	Mismatches:	22
Query Match:	35.28%	Indels:	1
DB:	11	Gaps:	1

US-09-896-522-2 (1-277) x US-09-918-995-23923 (1-455)

```
Qy      8 AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArgProPheLeuIle 26
      |||
      37 GACAGCGAGACGACCCCTGCGAAGACCACCGAGCCGCGGCGAGCCCTTCTTATATA 96
Qy      27 G1yValSerG1yG1yThraAlaSerG1yLysSerThraValCysG1uLysIleMetG1uLeu 46
      |||
```

Db 97 GGCCTCAGCGGGGGAACAGTAGCGGCAAGTCTCCGTGTGTGCTAAGATCGTCAGCTC 156
 Qy 47 LeuGlyGlnAenGluValGluGlnArgGlnArgValValIleLeuSerGlnAspArg 66
 Db 157 CTGGGGCAGATGAGTGAAGTACTATCGCAGAGAGGGTGTGATCTTGAAGCAGAGTACG 216
 Qy PheTyrIysValIleThrAlaGluGlnIysAlaIysAlaLeuIysGlyGlnTyrAspPhe 86
 Db 217 TTCTACCTGTCTTACTCTCGAGAGAGAGCCAAAGCCCTGMAAGGCCATTCTACTT 276
 Qy 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuIysAsnIleValGlu 106
 Db 277 GACCAACCGGATGCTTGTGACATGAACTCATTTCCAACACATCAAGAATACATCGAA 336
 Qy 107 GlyIysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGluThr 126
 Db 337 GGGAAAGATGCCAGATCCCGGTATGATCTTGTCTCCATCCCGAGAGAGAGACA 396
 Qy 127 ThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSer 145
 Db 397 GTTACTGTCTATCCCGCAGACGTGTGCTCTTTGAAGGATCTTGCCCTTCTACTCC 453

RESULT 14 US-09-070-927A-215/c

Sequence 215, Application US/09070927A
 Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunach
 Patrick J. Dillon
 Steven Barash
 TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
 NUMBER OF SEQUENCES: 982
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A
 FILING DATE: 04-May-2000
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655
 FILING DATE: 1997-05-16
 APPLICATION NUMBER: 60/044,031
 FILING DATE: 1997-05-06
 APPLICATION NUMBER: 60/066,009
 FILING DATE: 1997-11-14
 ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 215:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3662 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 215:
 US-09-070-927A-215

Alignment Scores:

Pred. No.: 1.64e-43 Length: 3662

Score: 427.00 Matches: 82
 Percent Similarity: 62.62% Conservative: 52
 Best Local Similarity: 38.32% Mismatches: 64
 Query Match: 29.45% Indels: 16
 DB: 10 Gaps: 4

US-09-896-522-2 (1-277) x US-09-070-927A-215 (1-3662)

Qy 22 ArgProPheLeuIleGlyValSerGlyThrAlaSerGlyPheSerThrValCysGlu 41
 Db 3328 CAGCCATTTCATCTGTGTACTGGGCGATCGAAGGCGGAAACCAAGTTAGCCCT 3269
 Qy 42 LysIleMetGluLeuLeuGlyGlnAenGluValGluGlnArgValValIle 61
 Db 3268 GCGATTTTAATAATTTTCCGATCAT-----TCAATTATGATG 3230
 Qy 62 LeuSerGlnAspArgPheTyrIys-----ValLeuThrAlaGluGlnValAlaIys 78
 Db 3229 TTGAACACGATCTTATATTAAGATCAGAGCCATTGTGATTGAAGAACCTTAAAT 3170
 Qy 79 AlaLeuIysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHis 98
 Db 3169 ACC-----AATTATGATCATCTTTTGGCTTGTATACAGATTTATTGATT 3125
 Qy 99 ArgThrLeuIysAsnIleValGluGlyThrValGluValProThrTyrAspPheVal 118
 Db 3124 CAACATGTTGAACACGTTAAATTTATCAAGCATGAAAGCCAGTTATGACTACGTT 3065
 Qy 119 ThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlu 138
 Db 3064 GCCCATCAAGAAAGTACAGAAACGATGATCAAGAACCGGAAAGATTAATTCTTGAA 3005
 Qy 139 GlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheVal 158
 Db 3004 GGGATTTTAATTATTAAGAAATCCCGCTTAAGGATTTATATGATTTAAAGGTATGTG 2945
 Qy 159 AspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal--ArgArgGly 177
 Db 2944 GATACGATGATGATCATCGTATTTATTCGATCAAGGTGATATGAAAGCGTGGC 2885
 Qy 178 ArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValAspProAlaPheGluGlu 197
 Db 2884 CGTACGTTGATTCGTGATTTGAACAATATTGACAGCTCCTCAAGCCCATGTATCAACA 2825
 Qy 198 PheCysLeuProThrIysIysTyrAlaAspValIleIleProArgGlyValAspAsnMet 217
 Db 2824 TTATTGACCTTACGAACGTTATGCAATATCATTTGTGCCAGAAAGTGGCGAAACACAC 2765
 Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeu 231
 Db 2764 GTCGCTATTGATTTAATTATACAAAGTGATGATTTTTA 2723

RESULT 15

US-10-329-960-1
 Sequence 1, Application US/10329960
 Publication No. US20030099277A1

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragment of the
 TITLE OF INVENTION: Thereof, and Uses Thereof

FILE REFERENCE: PB186P1

CURRENT APPLICATION NUMBER: US/10/329,960

CURRENT FILING DATE: 2003-01-02

PRIOR APPLICATION NUMBER: US 09/643,990

PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: US 08/487,429

PRIOR FILING DATE: 1995-06-07

PRIOR APPLICATION NUMBER: US 08/426,787

PRIOR FILING DATE: 1995-04-21

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 1830121

```
? TYPE: DNA
? ORGANISM: Haemophilus influenzae
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (4747)..(4747)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (9921)..(9921)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (10150)..(10150)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (29298)..(29298)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (36543)..(36543)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (36551)..(36551)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (36636)..(36636)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (40808)..(40808)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (44416)..(44416)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (44905)..(44905)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (44975)..(44975)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (45593)..(45593)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (45732)..(45732)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (47036)..(47036)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (51334)..(51334)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (51602)..(51602)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (51786)..(51786)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (51805)..(51805)
? OTHER INFORMATION: n equals a, t, g or c

? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (55369)..(55369)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (65309)..(65309)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (65313)..(65313)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (80024)..(80024)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (100091)..(100091)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (102696)..(102696)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (105121)..(105121)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (107248)..(107248)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (117136)..(117136)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (119750)..(119750)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (119924)..(119924)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (120038)..(120038)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (121344)..(121344)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (122167)..(122167)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (122336)..(122336)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (131340)..(131340)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (131360)..(131360)
? OTHER INFORMATION: n equals a, t, g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (139510)..(139510)
? OTHER INFORMATION: n equals a, t, g or c
```

```
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (140398) .. (140398)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (142750) .. (142750)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145058) .. (145058)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145171) .. (145171)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (145942) .. (145942)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147197) .. (147197)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (150841) .. (150841)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152500) .. (152500)
/ OTHER INFORMATION: n equals a, t, g or c
/
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (152530) .. (152530)
```

```
Alignment Scores:
Pred. No.: 6.76e-34 Length: 1830121
Score: 384.00 Matches: 80
Percent Similarity: 60.56% Conservative: 49
Best Local Similarity: 37.56% Mismatches: 70
Query Match: 26.48% Indels: 14
DB: 14 Gaps: 4
```

US-09-896-522-2 (1-277) x US-10-329-960-1 (1-1830121)

```
QY 25 LeuIleGlyValSerGlyThrAlaSerGlyLys-----SerThrValCys 40
DB 146027 ATCATCGGCATCACGCGTCATCTGCATCAGGCAGAAAGTTCTTATCGCTTCCACTGTTTCAC 146086
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
DB 146087 AAAGAGCTCTGTATGAAGCTAGCGTCCAGCAAGAAATGCT-----146125
QY 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
DB 146126 ATCATTTACTGAAGACAGTTACTACTATAAGATCAAAAGTCAATTAGAAATAGCTAGCGGGTA 146185
QY 81 LysGlyGlnIleLysPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 146186 AAA-----ACGAATTACGATCATCCAAACTGTAGTGCAGATTACTTATCCACCAT 146239
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
DB 146240 TTAATAAATCTAAAAAATGCGACAGTGCAGTATGCTGTTTATAGCTATGTAAACAT 146299
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
DB 146300 ACTCGCACCAACAGAAACAGCATTTACACCAAAACGATCGTAATTGTAGAAAGGATT 146359
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 146360 TTATTACTTACCGATGAAGAGTAGCGCAATTAGCCGATATTCTGTATTGTAGACACA 146419
```

```
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspVal--ArgArgIleArgAsp 179
DB 146420 CCACTTGATATTGTTGTTTCATCCGCCGTTTACACGATATGGAAGACGTGCTCT 146479
QY 180 LeuGluGlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGluPheCys 199
DB 146480 CTACAAATCAGTGAATTGATCAATATCGTCAACCGTGCATCAATGTTCTTACCAATTATT 146539
QY 200 LeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAla 219
DB 146540 GAGCCGCTTAACAAATATGCGGAAATTTGATTCCTCGTGGTGTAAATAATCCCATTTGCG 146599
QY 220 IleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn 232
DB 146600 ATCAATATGTTAAAGCTCAATCCCTTCAATTATGAAAT 146638
```

Search completed: November 25, 2003, 10:23:54
Job time : 917 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:09:36 ; Search time 317 Seconds

(without alignments)
2358.815 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450

Sequence: 1 MASAGEDESPAPADRRH.....HPGMLTSGKSHLESSRRH 277

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seque, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-Q=/c9n2_1/USPTO.spool/US09896522/runcat_21112003_184353_4997/apd_query.fasta_1.455
-DB=N Geneseq_19Jun03 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896522.OCGN_1.1.312.0/runcat_21112003_184353_4997 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_19Jun03:*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1450	100.0	834	23	AAK98735	DNA of a human uri
2	1450	100.0	1288	22	AA158658	Human polynucleoti
3	1450	100.0	1624	24	AA271786	Human 57658 DNA.
4	1450	100.0	2152	22	AA160444	Human polynucleoti
5	1450	100.0	2160	22	AA115621	Human cDNA sequenc
6	1380	95.2	1396	24	AB189762	Human polynucleoti
7	1372.5	94.7	900	22	AAH75355	Human uridine kina
8	1074.5	74.1	753	22	AAH04832	Human cDNA clone (
9	941.5	64.9	1322	22	AAH23801	Human transferase
10	940	64.8	1402	22	AA157850	Human polynucleoti
11	924	63.7	734	24	AB551801	Human mdct cDNA in
12	907.5	62.6	1322	24	AB054470	Human ovarian anti
13	902.5	62.2	1310	21	AAH15785	Human prostate can
14	794.5	54.8	1316	22	AA159636	Human polynucleoti
15	738	50.9	822	23	AB110981	Drosophila melanog
16	736.5	50.8	549	25	ABT33606	Anticancer agent t
17	701.5	48.4	715	20	AA215454	Human gene express
18	647	44.6	19125	22	AA136440	Human musculoskele
19	647	44.6	19125	25	ABX59428	CDNA encoding nove
20	622.5	42.9	3413	23	AB110980	Drosophila melanog
21	590	40.7	980	23	AA591403	DNA encoding novel
22	578	39.9	1770	21	AAAC4047	Zea mays DNA fragm
23	570.5	39.3	1536	21	AAAC48913	Arabidopsis thalia
24	561.5	38.7	1940	21	AAAC3670	Zea mays DNA fragm
25	545.5	37.6	1647	24	ABK90806	CDNA encoding huma
26	544	37.5	1401	24	AB214138	Arabidopsis thalia
27	542	37.4	1856	23	AB120375	Drosophila melanog
28	542	37.4	2319	23	AB120375	Drosophila melanog
29	542	37.4	1813	23	AB108329	Arabidopsis thalia
30	540.5	37.3	3193	24	ABK59846	Novel human coding
31	540.5	37.3	1959	22	AA545081	CDNA encoding nove
32	529	36.5	1326	21	AAAC5084	Arabidopsis thalia
33	529	36.5	1831	21	AAAC37084	Arabidopsis thalia
34	526.5	36.3	2065	21	AAAC75866	Human ORF1421
35	518	35.7	1503	21	AAAC48458	Arabidopsis thalia
36	507.5	35.0	1500	21	AAAC49434	Arabidopsis thalia
37	507.5	35.0	1703	21	AAAC42628	Arabidopsis thalia
38	490.5	33.8	1838	21	AAAC51009	Arabidopsis thalia
39	485.5	33.5	3881	23	ABV23398	Human prostate exp
40	485.5	33.5	3881	23	ABV24102	Human prostate exp
41	485.5	33.5	3881	23	ABV24281	Human prostate exp
42	485.5	33.5	3881	23	ABV29881	Human prostate exp
43	485.5	33.5	3881	23	ABV29883	Human prostate exp
44	485.5	33.5	3881	23	ABV30091	Human prostate exp
45	473.5	32.7	7248	23	AB108328	Drosophila melanog

ALIGNMENTS

RESULT 1
AAK98735
ID AAK98735 standard; DNA; 834 BP.
XX AAK98735,
XX
XX
XX 02-MAY-2002 (first entry)
XX
XX
XX DNA of a human uridine kinase (UDK).
XX
XX
XX Human; uridine kinase; diagnostic assay; mutation detection; UDK.
XX
XX prob; chromosome localisation study; tissue expression; gene therapy;
XX antibody; vaccine; human ovarian cancer; immunological disorder;
XX human colon carcinoma; immunogen; ds.
XX
XX Homo sapiens.
XX OS

Key	Location/Qualifiers
FT CDS	1..780
FT	/+tag= a
FT	/partial
FT	/note= "No stop codon"
PX	
PN	WO200172963-A2.
PD	
XX	
PD	04-OCT-2001.
XX	
PF	27-MAR-2001; 2001WO-US09663.
XX	
PR	27-MAR-2000; 2000US-0536647.
XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.
XX	
P1	Ho YS, Johnson RK;
XX	
DR	WPI: 2001-626259/72.
DR	P-PSDB; AAO14412.
XX	
PT	Novel human uridine kinase polypeptides useful for treating cancers,
PT	and to identify agonists and antagonists of the polypeptide useful for
PT	treating conditions associated with uridine kinase imbalance -
XX	
PS	Claim 5; Page 22-23; 31pp: English.
XX	
CC	The invention relates to newly identified human uridine kinase (UDK)
CC	polypeptides and polynucleotides and methods for producing such
CC	polypeptides by recombinant techniques. Also disclosed in the invention
CC	are methods for utilizing uridine kinase polypeptides and polynucleotides
CC	in diagnostic assays. The polynucleotides and polypeptides of the
CC	invention may be used as diagnostic reagents by detecting mutations in an
CC	associated gene. An array of oligonucleotide probes comprising the
CC	uridine kinase polynucleotide sequence or fragments thereof can be
CC	constructed to conduct efficient screening of genetic mutations, for
CC	example. Detection of abnormally decreased or increased levels of
CC	polypeptide or mRNA expression may also be used for diagnosing or
CC	determining susceptibility of a subject to a disease of the invention.
CC	The polynucleotide sequences of the invention can be used for chromosome
CC	localisation studies and tissue expression studies. The polypeptides of
CC	the invention or fragments thereof may be used as immunogens to produce
CC	antibodies. These antibodies may be employed to isolate or identify
CC	clones expressing the polypeptide. The polypeptides and polynucleotides
CC	of the invention can be used as a vaccine or in gene therapy to treat
CC	diseases such as human ovarian cancer, human colon carcinomas, and
CC	immunological disorders. This polynucleotide sequence represents the DNA
CC	of a human uridine kinase of the invention.
XX	
SQ	Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;
Alignment Scores:	
Pred. No.:	9.71e-154 Length: 834
Score:	1450.00 Matches: 277
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	23 Gaps: 0
US-09-896-522-2 (1-277) x AAK38735 (1-834)	
OY	1 MetAlaSerAlaGlyGlyGluAbpCYsGUsErProAlaProGluAlaAspArgProHis 20
Dd	1 ATGGCTTCGGCGGGAGGCAGACTGCCAGAACCCCCCGCGCGAGGCGCAGACCGTCCGCAC 60
OY	21 GluNarGPrOheUeuIleGlyValSerGlyGlyThrAlaSerGlyLYVSeTerThValCys 40
Dd	61 CAGCGGCCCTTCTCGATAGGGGTGAACGGCGCACCTGCAAGGGGAAAGTCAACCGTGTT 120
OY	41 GluLysGIleMetGluLeuLeuGlyGlnNaSGlValGluGlnGlnGlnGlnGlnGlnGlnGln 60
Dd	121 GAGAAGATCATGAGATTGCTGGGACACGAGGTGAAAACAACGGCACGGGAAAGGTGTC 180

QY	61	IIleuSeGclAAspRgRheTyTysValleuThrIaGluGlnYsAlaYsAlaLeu	80
Db	181	ATCTGAAGCCAGGACACAGGTTCTTCAAGGCTCTGAAGCGAGGACGAAGGCCAAGGCTTG	240
QY	81	LyGclYGlntYrAsnPhaAspHiAsProAlaAlaPhaAspAenAspLeuMetHisArgThr	100
Db	241	AAAGGACAGTACAAATTTTGTACCATCCAAATGCCCTTTGACAAATGATTTGAATGACACAGCACT	300
QY	101	LeuLYAsAnIleValGluGlySerThrValGluValProThrTYrAspPheValThrHis	120
Db	301	CTGAAGACATCTGTGAGGAGGACAAACGGTGAAGGTGCCGACCTTATGATTTTGTGACACAC	360
QY	121	SeArIgleuProGluThrThrValValTYrProAlaAspValValleuPheGluGlyIle	140
Db	361	TCAAGGTTACAGAGACCAAGGTGGTCTTACCTGGCGGACGTGGTCTTGTGTTGAGGCGATC	420
QY	141	LeuValPheTySerGlnGlnIuIleArgMetPheHisIleuArgLeuPheValAspThr	160
Db	421	TTGATGTCTTACGCGCAGGAGATCCGGGACATGTTCACCTGGCGCTTCTGTGTTGAGACAC	480
QY	161	AspSerAspValArgLeuSerArgArgValleuArgAspValArgArgGlyArgAspLeu	180
Db	481	GACTCCGACCTCAGCGCTGTCTGAAAGATTCCTCCGGGACGTGGCCGACGAGGAGGACCTG	540
QY	181	GluGlnIleleuThrGlnTYrThrThrPheValLYsProAlaPheGluGluPheCysLeu	200
Db	541	GAGCAGATCTTGCACGAGTACACCACTTCGTGGAAAGCCGGCGCTTCGAGAGTTCTGCTG	600
QY	201	ProThrLYsTYrTYrAlaAspValIleIleProArgGlyValAspAenMetValAlaIle	220
Db	601	CCGACAAAGAGATGCCCATGTGATTCATCCACGAGAGTGGCAATATGCTTGCCATC	660
QY	221	AsnleuIleValGlnHisIleGlnAspIleleuAenGlyAspIleCysLYsTPrHisArg	240
Db	661	AACCTGATCGTGACACATCCAGGACATTCGATGGTGTGACATCTCCAATGGCACCGA	720
QY	241	GlyGlySerAsnGlyArgSerTYrLYsArgThrPheSerGluPurProGlyAspHisArgGly	260
Db	721	GGAAGGTCATAGGAGGAGCTTACAAAGCGGACCTTTCTGAGCCAGGAGGACACCTTGAG	780
QY	261	MetleuThrSerGlyLYsArgSerHisIleuGluIleuSerSerArgProHis	277
Db	781	ATGCTGACCTCTGGCAACGATCCATTTGGAGTCCACGACAGACCCCAAC	831
RESULT 2			
AA158658	ID	AA158658 standard; cDNA; 1288 BP.	
XX	AA158658;		
XX	22-OCT-2001 (first entry)		
DE	Human polynucleotide SEQ ID NO 861.		
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukemia; ss.		
OS	Homo sapiens.		
XX	WO200153312-A1.		
PM	26-JUL-2001.		
XX	26-DEC-2000; 2000MO-US34263.		
PF	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		

PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
 PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM39502.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1; SEQ ID NO 861; 10078BP; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1,796-153 Length: 1288
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-896-522-2 (1-277) x AA158658 (1-1288)
 QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGCGCCGCGCCAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 155 CACGGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTCCAGCGGGAAGTCGACGGTGTG 214
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 DB 215 GAGAAGATCATGAGTGTGCTGGGACGAAACAGGTGGAACAGCGGCGACGGAAGGTGCTC 274
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 275 ATCTTAGCCAGACAGAGCTTCTACAGAGTCTCTGACGACAGACGAAAGGCCAGCTTG 334
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 335 AAAGGACAGTACATTTTGCACATCCAGATGCCCTTTATATGATTTTGGATGCACAGGACT 394
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 395 CTGAGAGAACATCGTGGGCGCAAAACGCTGAGAGTGGCGACTATGATTTTGTGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

DB 455 TCAGAGTTACAGAGACACACGGGTGCTACCCCTGGCGGAGTGGTTCTGTTAGAGGCATC 514
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 515 TTGGTGTCTACAGCAGGAGAGATCCCGGACATGTTCCACTGCGCTCTTCTGGACACC 574
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 DB 575 GACTCCGAGTCAGGCTGCTCGAAGAGTTCCTCGGAGCGTGCGCGAGGAGGACCTG 634
 QY 181 GluGlnIleuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 635 GAGCAGATTCTGACGAGATACACACCTTCGTGAACCGGCTTCGAGGAGTCTGCTC 694
 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAsnMetValAlaIle 220
 DB 695 CCGACAAAGAGATATCCGATGATCATCCACGAGAGTGGACAATATGATTGGCCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheHisArg 240
 DB 755 AACCTGATGTGAGACATCCAGACATTCGAATGTGACATTCGCAATGGCACCGA 814
 QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
 DB 815 GAGGGTCCAAATGGCGGAGCTACAAAGCCGACTTTCTGAGCCAGGAGGACCACTTGG 874
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 875 ATGCTGACCTCTGGCAACGATCACTTTGAGTCCAGACAGACACCCAC 925
 RESULT 3
 AAD27186
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 AC AAD27186;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 KW Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticovulcent; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH CDS
 FT 94..927 Location/Qualifiers
 FT /*tag= a
 FT /product= "Human 57658 protein"
 FT /*tag= b
 FT /note= "This region is specifically claimed as
 FT SEQ ID NO: 3 in claim 1 of the specification"
 XX
 PN WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;

CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, Leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

CC Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Alignment Scores:

Pred. No.:	3,68e-153	Length:	2152
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
	22	Gaps:	0

US-09-896-522-2 (1-277) x AA160444 (1-2152)

QY 1 MetLAserrAlaGlyGlyLysAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 59 ATGGCTTCGGCGGGAGCGAAGACTGGAGAGCCCGCGCGGCGCGCCGTCACAC 118
 QY 21 GlnArgProPheLeuIleGlyValSerGlyValThrAlaSerGlyLysSerThrValCys 40
 DB 119 CAGGGGCCCTTCCTGATNAGGGGTAGCGCGCGCACTCCAGCGGGAAGTGCACCTGTGT 178
 QY 41 GlnLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 DB 179 GAGAGATCATGAGTGTCTGGACAGAACGAGGTGAAACGCGGACCGGAAGTGTGTC 238
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysValAlaLysLeu 80
 DB 239 ATCTGTGCGCAGGACGAGTCTTCAAGAGTCTGACGCGACGAGGCGCAAGCCCTTG 298
 QY 81 LysGlyGlnTyrAsnProAspHisProAspAlaPheAspAspLeuMetHisArgThr 100
 DB 299 AAGAGCAGTACATTTTGCATCCAGATGCTTTATTAATGATTTGATGACACAGACT 358
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 359 CTGAAAGAACTCGTGGAGGGCAAAACGTTGAGGTGCGACCTATGATTTTGTGACACAC 418
 QY 121 SerArgLeuProGluThrThrValValValTyrProAlaAspValValLeuPheGluGlyIle 140
 DB 419 TCAGAGTTTACACAGACACGCGTGTCTACCTCGGACGCGGTCTGTGTGAGGCGATC 478
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
 DB 479 TTGGTGTCTTACACCGACGAGATCCGGAGCATGTTCACCTCGCCCTTCTGTGAGACAC 538
 QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 539 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGACGTCGCGCGGAGGAGGACCTG 598
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 599 GAGAGATTCTGACGACGTACACACACTTGTGTAAGCCGCGCTTCGAGGAGTCTGCTG 658
 QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
 DB 659 CCGAGCAAGAAGTATGCCGATGTATATCCCGGAGAGATGAGCAATATGTTGCCATC 718
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
 DB 719 AACCTGATCTGTGACGACATCCAGGACATCTGATGTGATCTCTCAATATGGCACCGA 778

QY 241 GlyGlySerAsnGlyValArgSerTyrLysArgThrPheSerGluProGluAspHisProGly 260
 DB 779 GAGAGGCTCATGAGGAGACTCAAGCGGACCTTTTCTGAGCAAGGAGACCCCTGGG 838
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 839 ATCTGACCTCTTGGCAACGCGTCACTTTGGATCTCAGACGACAGACCCAC 889
 RESULT 5
 AAH15621
 ID AAH15621 standard; cDNA; 2160 BP.
 AC AAH15621;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:13951.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 XX
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JUN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH1632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

Alignment Scores:

Pred. No.:	3.7e-153	Length:	2160
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
		Gaps:	0

US-09-896-522-2 (1-277) x ABL15621 (1-2160)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
   |||||
Db 95 ATGGCTTCGGCGGAGCGGAGAGCTGCGAGAGCCCGCGGAGGCGGAGCCGTCGCGAC 154
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
   |||||
Db 155 CAGCGGCCCTTCGATAGGGGTGAGCGCGGACCTGCGACCGGGAAAGTCGACCGTGTGT 214
QY 41 GluLysIleMetGluLeuLeuGlyGluAsnGluValGluGluAsnGluArgValVal 60
   |||||
Db 215 GAGAAAGATCATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACCGGAAAGTGTCT 274
QY 61 IleLeuSerGluAspArgPheThrValLeuThrAlaGluGluValAlaValAlaLeu 80
   |||||
Db 275 ATCTGACCGACGAGACAGGTTCTACAAAGGTCCTGACGCGACAGACAGAAAGCCGCTTG 334
QY 81 LysGlyGlnIleValAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
   |||||
Db 335 AAAGACAGTACAAATTTGACCATTCAGATGCTTTGATTAATGATTGATCAGACGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrValAspPheValThrHis 120
   |||||
Db 395 CTGAAGAACAATCTGTCGAGGCGAAACGCTGAGAGGCCGACCTATGATTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValIleProAlaAspValValLeuPheGluGlyIle 140
   |||||
Db 455 TCAAGGTACACAGAGACACGCGTGTCTACCTCGGACGCGTGTCTGTGAGGAGCATC 514
QY 141 LeuValPheTyrSerGlnGluIleLeuArgMetPheHisLeuValLeuPheValAspThr 160
   |||||
Db 515 TTGGTGTCTACGACGAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACCC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
   |||||
Db 575 GACTCCGACGTCAGGCTGTCTCGAAGACTTCTCGGAGACGTCGCCGAGGAGGAGACTG 634
QY 181 GluGlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGluPheCysLeu 200
   |||||
Db 635 GAGCAGAGATTGACGCGACGTACACCACTTCGTGAAGCCGCGCTTCGAGAGAGTTCGCTG 694
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
   |||||
Db 695 CCGACAAAGAGATGCGGATGTATCCACAGAGAGTGAGCAATATGTTGGCCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
   |||||
Db 755 AACCTGATCGTCGACGACATCCAGGACATTTCTGAATGTGACATCTGCAATATGGACCGA 814
QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
   |||||
Db 815 GAGAGGTCTCAATGGCGGAGCTACCAAGCGGACCTTTTCTGAGCAGGGGACCACTTGGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisIleGluLysSerSerArgProHis 277
   |||||
Db 875 ATGCTGACCTCTGGCAACGGTCAATTGAGTTCACAGACAGACCCAC 925

```

RESULT 6
ABL89762
ID ABL89762 standard: cDNA, 1396 BP.

XX ABL89762,
XX AC

DT 24-MAY-2002 (first entry)

XX DE Human polynucleotide SEQ ID NO 324.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 XX KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 XX KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX KW neurological disease; infection; human; secreted protein; gene; ss.

XX OS Homo sapiens.

XX PN WO200190304-A2.

XX PD 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX PR 19-MAY-2000; 2000US-205515P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX XX WPI: 2002-122018/16.

XX DR P-PSDB; ABB89353.

XX PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 XX PT prevention of neural, immune system, muscular, reproductive,
 XX PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 XX PT disorders -

XX PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.

XX CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 XX CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 XX CC medical conditions e.g. by protein or gene therapy. The genes are
 XX CC isolated from a range of human tissues disclosed in the specification.
 XX CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 XX CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 XX CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 XX CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 XX CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 XX CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 XX CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 XX CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 XX CC (d) wound healing; (e) neurological diseases e.g. cerebral ischaemia;
 XX CC and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 XX CC and parasitic infections.
 XX CC Note: The sequence data for this patent did not form part of the
 XX CC printed specification, but was obtained in electronic format directly
 XX CC from WIPO at http://wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;

Alignment Scores:

Pred. No.:	1.56e-145	Length:	1396
Score:	1380.00 <td>Matches:</td> <td>272</td>	Matches:	272
Percent Similarity:	98.19%	Conservative:	0
Best Local Similarity:	98.19%	Mismatches:	5
Query Match:	95.17%	Indels:	2
		Gaps:	0

US-09-896-522-2 (1-277) x ABL89762 (1-1396)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
   |||||
Db 41 ATGGCTTCGGCGGAGCGGAGAGCTGCGAGAGCCCGCGGAGGCGGAGCCGTCGCGAC 100
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
   |||||
Db 101 CAGCGGCCCTTCGATAGGGGTGAGCGCGGACCTGCGACCGGGAAAGTCGACCGTGTCT 160

```

QY 41 GlnuyllemerGluLeuLeuGlnAsnGlnValGlnGlnArgGlnArgValVal 60
DB 161 GAGAGATCTAGAGTTGCTGGGACAGACAGAGTGGAAACCGGACGGAGGTGTC 220
QY 61 IleuSerGlnAspArgPheTyrLysValIleuThrAglGlnIlySalAluValIleu 80
DB 221 ATCTGAGCAGGACGAGGTTCTCAAGGTCTTGACGAGGACGAGGACGAGGCTTG 280
QY 81 LysGlnIlyLysAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 281 AAGAGACAGTACAAATTTTGAACATCCAGATGCTTTGATATATGATTTGATGACAGACT 340
QY 101 LeuLysAsnIleValGlnGlnIlyLysThrValGlnValProThrTyrAspPheValThHis 120
DB 341 CTGAGAGACATCTGAGAGGCGAAACGCTGGAGGTGCGACCTTGATTTTGACACAC 400
QY 121 SerArgLeuProGlnThrThrValValIlyProAlaAspValValLeuPheGlnIlyIle 140
DB 401 TCAAGGTTACAGAGACAGGAGTGTCTACCTCGGACGAGGTTCTGTTGAGGACATC 460
QY 141 LeuValPheTyrSerGlnIlyLysArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 461 TTGGTGTCTACAGCCGAGAGATCCGAGACATGTTCCACCTGCGCTTCTGAGACACC 520
QY 161 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlnIlyArgAspLeu 180
DB 521 GACTCCGACGTCAGGCTGTCTCGAAGAKTTCTCCGAGACGTGGCCGAGGAGGAGACTG 580
QY 181 GlnGlnIleLeuThrGlnIlyThrThrPheValLysProAlaPheGlnIlyPheCysLeu 200
DB 581 GAGCAGATTCGACGCA-TACACACACTTGTGTAAGCCGCGCTTCGAGAGTTCTGCTG 639
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlnValAlaAsnMetValAlaIle 220
DB 640 CCGACAAAGAGATATGCGATGATATCCACAGAGAGTGAACAATATGTTGCCATC 699
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnIlyAspIleCysLysThrPheArg 240
DB 700 AACCTGATCTGACGACATCCAGACATTTGATGATGATGATGATGATGATGATGATG 759
QY 241 GlnGlnIlySerAsnGlnIlyArgSerTyrLysArgThrPheSerGlnProGlnIlyAspHisProGln 260
DB 760 GAGAGGTCCAAATGGGCGAG-TACAAAGCGACCTTTCTTACGAGGAGGACACCTGGG 818
QY 261 MetLeuThrSerGlnIlyLysArgSerHisLeuGlnIlySerSerArgProHis 277
DB 819 ATCTGACCTCTGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATG 869
RESULT 7
AAH75355 ID AAH75355 standard; cDNA; 900 BP.
AC AAH75355;
DT 02-OCT-2001 (first entry)
DE Human uridine kinase encoding cDNA.
KW Human; uridine kinase; UK; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 8..838
FT /tag= a
FT /product= "uridine kinase"
FT /note= "claimed in claim 1"
XX
XX CNI287172-A.
XX 14-MAR-2001.
XX
XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.
PR (UYFU-) UNIV FUDAN.
PA Yu L, Zhao Y, Zhang H;
PI WPL; 2001-409529/44.
DR P-PSDB; AAG64506.
XX
XX Human uridine kinase and its coding sequence, preparation and
PT application -
XX
PS Claim 1; Page 15(Disclosure); 20pp; Chinese.
CC The invention relates to human uridine kinase (UK).
XX
SQ Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
Alignment Scores:
Pred. No.: 5,92e-145 Length: 900
Score: 1372.50 Matches: 270
Percent Similarity: 98.19% Conservative: 2
Best Local Similarity: 97.47% Mismatches: 5
Query Match: 94.66% Indels: 3
DB: 22 Gaps: 0
US-09-896-522-2 (1-277) x AAH75355 (1-900)
QY 1 MetAlaSerAlaGlnIlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
DB 8 ATGGCTTCGGCGGAGGCGAAGG-TCCGAGAG-CCGCGCCGAGAGG-AAACCTCCGAC 64
QY 21 GlnArgProPheLeuIleGlnIlySerGlnIlyThrAlaSerGlnIlySerThrValCys 40
DB 65 CAGCGTCCCTTCTGATAGGAGGTGAGCGGCGGACATCCAGCGGAAAGTGCACCTGTGT 124
QY 41 GlnuyllemerGluLeuLeuGlnIlyAsnGlnValGlnGlnArgGlnArgValVal 60
DB 125 GAGAGATCATGAGTTGCTGGGACAGAACAGAGTGAACAGCGGACCGGTAAGTGTGTC 184
QY 61 IleuSerGlnAspArgPheTyrLysValIleuThrAglGlnIlySalAluValIleu 80
DB 185 ATCTGAGCAGGACGAGTCTCAAGGTCTTGACGAGAGGACGAGGACGAGGCTTG 244
QY 81 LysGlnIlyLysAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 245 AAGAGACAGTACAAATTTTGAACATCCAGATGCTTTGATGATGATGATGATGATGATG 304
QY 101 LeuLysAsnIleValGlnGlnIlyLysThrValGlnValProThrTyrAspPheValThHis 120
DB 305 CTGAAAGACATCTGAGAGGCGAAACGAGTGAAGTGCACCTATATATTTTGACACAC 364
QY 121 SerArgLeuProGlnThrThrValValIlyProAlaAspValValLeuPheGlnIlyIle 140
DB 365 TCAAGGTTACAGAGACAGGAGTGTCTACCTCGGAGAGTGTGTTTGAGGACATC 424
QY 141 LeuValPheTyrSerGlnIlyLysArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 425 TTGGTGTCTACAGCAGAGATCCGAGACATGTTCCACTGGGCTCTTCTGAGACACC 484
QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlnIlyArgAspLeu 180
DB 485 GACTCCGAGGTGAGTGTCTGAAAGATTTCTCCGAGAGGTGGCGCGGAGGAGACTG 544
QY 181 GlnGlnIleLeuThrGlnIlyThrThrPheValLysProAlaPheGlnIlyPheCysLeu 200
DB 545 GAGCAGATTCGAGAGGACGACACACCTTCGTAAGCCGCGCTTCAGAGAGTTCTGCTG 604
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlnValAlaAsnMetValAlaIle 220
DB 605 CCGACAAAGAGATATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 664

QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 Db 665 AACCTGATCGTCGACGACATCCAGGACATTCTGATGTGACATCTGCAAAATGGCACC 724
 QY 241 GlyGlySerAsnGlyArgSerTyrIlysaArgThrPheSerGluProGlyAspHisProGly 260
 Db 725 GGAGGGTTCACATGGGCGGACGACGACGACCTTTCTGAGCCAGGGGACCCACTGGG 784
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 Db 785 ATGCTGACCTCTGGCAACGGTCACTTGGAGTCCAGCAGACCCAC 835
 RESULT 8
 AAH04832
 ID AAH04832 standard; cDNA; 753 BP.
 XX
 AC AAH04832;
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:1667.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 XX Claim 1; SEQ ID 1667; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination
 CC of the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.
 XX
 SQ Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other:
 Alignment Scores:
 Pred. No.: 1,83e-111 Length: 753
 Score: 1074.50 Matches: 213
 Percent Similarity: 97.26% Conserves: 0
 Best Local Similarity: 97.26% Mismatches: 3
 Query Match: 74.10% Indels: 3
 DB: 22 Gaps: 2
 US-09-896-522-2 (1-277) x AAH04832 (1-753)
 QY 1 MetAlaSerAlaGlyGlyAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 95 ATGGCTTCGGGCGGAGCGCAAGACTGGAGAGCCCGCGGAGGCGGACCGCTCGCAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 155 CAGCGGCCCTTCTGTGATAGGGGTGAGCGGCGGACCTGCCAGCGGGAAGTGAACCTGTGT 214
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 Db 215 GAGAGATCATGGAGTGTGCTGGGACAGAAACGAGTGGAAACCGCGACGCGAGCTGTCTC 274
 QY 61 IleLeuSerGlnAspArgPheTyrLysValIleThrAlaGluGlnLysAlaLysValLeu 80
 Db 275 ATCCTGAGCCAGACAGGTCTCAAGGTCTTACGCGCAGAGCAAGAGCCAGGCTTGG 334
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 335 AAGGACAGTACATATTTTGACCATCCAGATGCTTGTATATATTTGATGACAGAGACT 394
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 395 CTGAGAGACATCGTGGAGGGGCAAAACGCTGAGAGTCCGCACTATGATTTTGTGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 455 TCAAGGTTCACAGAGACCGGTGTCTACCTCGGACGCTGTCTGTTGGGGCATC 514
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 515 TTGGTGTTCACAGCCAGAGATCCGGACATTTCCACTGCGCTCTTCGTGGACACC 574
 QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgLysArgLeu 180
 Db 575 GACTCCGACCTCAGGCTGTCTCCAGAGATTTCTCCGGAGCGTGGCCGAGGAGGACCTG 634
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 635 GAGCAGATTCTGACGCGATACACACCTTCGTGAACCCGGCTTCAGAGAGTTCC---TTG 691
 QY 201 ProThr-----LysLysTyrAlaAspValIleIleProArgGlyValAspAsnMet 217
 Db 692 CTTTGCACAAAGAAAGTATGCCGATGATCATTTCCACAGAGAGTGGCAATATAG 748
 RESULT 9
 AAH23801
 ID AAH23801 standard; cDNA; 1322 BP.
 XX
 AC AAH23801;
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transferrase HTFS-1 cDNA, SEQ ID NO:43.
 XX
 KM Human transferrase; HTFS; agonist; antagonist; cellular signalling;
 KM proliferation; cell proliferative disorder; immune disorder;
 KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
 KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
 KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;

KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
 KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
 KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
 KM gene therapy; drug screening; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 250..1025
 FT /tag= a
 FT /product= "HTFS-1"
 FT /function= "Transferase"
 XX
 XX MO200132888-A2.
 XX
 PD 10-MAY-2001.
 XX
 PF 02-NOV-2000; 2000MO-US30485.
 XX
 PR 04-NOV-1999; 99US-0163595.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
 PI Shih LL, Azimzai Y, Lu DAM, Baughn MR;
 XX
 DR WPI: 2001-328796/34.
 DR P-PSDB; AAB73494.
 XX
 XX Human transferase polypeptides and polynucleotides useful for
 PT diagnosis, prevention and treatment of cell proliferative and immune
 PT system disorders and for identifying agonists and antagonists -
 XX
 PS Claim 5; Page 135-136; 157pp; English.
 XX
 XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
 CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
 CC encoding them. The proteins play important roles in the regulation of
 CC cellular signalling and proliferation. The HTFS proteins are useful for
 CC screening compounds for their effectiveness as agonists or antagonists of
 CC transferase activity, or for compounds that specifically bind to an HTFS
 CC protein or which modulates the activity of an HTFS protein.
 CC Pharmaceutical compositions comprising an HTFS protein, HTFS
 CC agonist or antagonist, or genetic construct encoding an HTFS
 CC protein are useful for treating a disease or condition associated
 CC with decreased or increased expression of functional HTFS. Disorders
 CC which may be treated using such compositions include cell proliferative
 CC disorders and immune disorders. For example, diseases which may be
 CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
 CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
 CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
 CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
 CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
 CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
 CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
 CC transgenic animals to model human diseases, for diagnostic purposes and
 CC to generate hybridisation probes useful in mapping the naturally
 CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
 CC fragments are useful for screening libraries of compounds in a variety of
 CC drug screening techniques. Antibodies which specifically bind HTFS may be
 CC used for the diagnosis of disorders associated with the expression of
 CC HTFS, or in assays to monitor patients being treated with HTFS or
 CC agonists, antagonists or inhibitors of HTFS. The present sequence
 CC represents an HTFS protein-encoding cDNA of the invention.
 CC
 XX
 SO Sequence 1322 BP; 339 A; 359 C; 358 G; 266 T; 0 other;

Alignment Scores:
 Pred. No.:
 Score:
 Percent Similarity:

4e-96
 941.50
 78.87%

Length:
 Matches:
 Conservative:

1332
 192
 32

Best Local Similarity: 67.61% Mismatches: 37
 Query Match: 64.93% Indels: 23
 DB: 22 Gaps: 5
 US-09-896-522-2 (1-277) x AAH23801 (1-1322)
 QY 3 SerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgPro----- 19
 DB 223 GCACGGCGAGAGAGGGCGCGGACCAACAGCGCGGAGACGAGACAGCCGCGAAGC 282
 QY 20 HsGlnArg-----ProPheLeuIleGlyValSerGlyGlyThraLaser 34
 DB 283 CACACAGACGCCCAACAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGGAG 342
 QY 35 GlyLysSerThraValCysGluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGln 54
 DB 343 GCGAAGTCCTTCGTGTGTCTAAGATCGGACGCTCTGGGCGAGATAGAGGTAGCTAT 402
 QY 55 ArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlu 74
 DB 403 CGCCAGACAGAGGTGTCATCTGACCGAGATAGCTTACCGTCTCTTACCTCGAG 462
 QY 75 GlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsn 94
 DB 463 CAGAAAGCCAAAGCCCTGAAGGGCGAGTCACTTGAACCAACCGGATGCTTGAACAAT 522
 QY 95 AspleuMechHisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThr 114
 DB 523 GAACCTATCTCAAAACACTCAAAAGAAATCACTGAAGGAGAAACACTCAAGATCCCGTG 582
 QY 115 TyrAspPheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspVal 134
 DB 583 TATGACTTGTCTCCATTCCTCCGGAAGAGAGACAGTACTGTCTATCCCGACGCTG 642
 QY 135 ValLeuPheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeu 154
 DB 643 GGGCTTGTGAAGAGGATCTGGCGCTTCTACTCCAGAGAGTACAGAGACCTGTCCAGATG 702
 QY 155 ArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
 DB 703 AAGCTTTTGTGAATACATGATCGGACCGCGCTCTCAGGAGATATTAAAGGACATC 762
 QY 175 ---ArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysPro 193
 DB 763 AGCGAGAGGCGAGGAGATTTGACAGATTTTATCTCAATTAATTCGTTCAAGCT 822
 QY 194 AlaPheGluGluPheCysLeuProThrLysLysTyrAlaAspValIleLeuProArgGly 213
 DB 823 GCGTTTGAAGAAATTCGCTTGCACAAAGAAATGATGCTGATCATCCCTAGAGCT 882
 QY 214 ValAspAspMetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGly 233
 DB 883 GCGAGTAAATCTGTGGCCCTCAACCTCACTCGAGCAGCATCAGAGCACTCTGATATGA 942
 QY 234 AspIleCysLysTyrPheHisArgGlyLysSerAsnGlyArgSerTyrLysArgThrPheSer 253
 DB 943 -----GGGCGCTTCAAAAGGAGAC----- 963
 QY 254 GluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSer 273
 DB 964 ---AATGGTGTCTCAAGCGCTACACCCCTTACCGAAGAGGACAGGCATCGAGTCCAGC 1020
 QY 274 SerArgProHis 277
 DB 1021 AGCAGGCCCGCAT 1032
 RESULT 10
 AA157850 standard; cDNA; 1402 BP.
 ID AA157850
 XX
 AC AA157850;
 XX
 DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 53.
 XX XX
 KW Human; nootropic; immunosuppressant; cyrostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US34263.
 XX PR 21-JAN-2000; 2000US-0488725.
 XX PR 25-APR-2000; 2000US-0552317.
 XX PR 09-JUL-2000; 2000US-0598042.
 XX PR 19-JUL-2000; 2000US-0620312.
 XX PR 03-AUG-2000; 2000US-0653450.
 XX PR 14-SEP-2000; 2000US-0662191.
 XX PR 19-OCT-2000; 2000US-0693036.
 XX PR 29-NOV-2000; 2000US-0727344.
 XX PA (HSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AD, Yang Y, Zhang J,
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-4442253/47.
 DR P-PSDB; AAM38694.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS Claim 1; SEQ ID NO 53; 10078pp; English.
 XX CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, Leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX SQ Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;
 Alignment Scores:
 Pred. No.: 6.42e-96 Length: 1402
 Score: 940.00 Matches: 194
 Percent Similarity: 74.66% Conservative: 27
 Best Local Similarity: 65.54% Mismatches: 41
 Query Match: 64.83% Indels: 34
 DB: 22 Gaps: 5
 US-09-896-522-2 (1-277) x AA157850 (1-1402)
 QY 2 Alasertaglygylgylu-----
 DB 227 GCGGGCGCGCGCGCGCGCTGCGTTCGACACGACGCGGAGAGGCGCGCGCA 286

QY 8 -----AspCysGluSerProAlaProGluAlaAspArgProHis---GlnArg 22
 DB 287 ACCATGCGCGGAGACGACGACGACCGCTGCAGAACACCGACGACGCGCGCGAG 346
 QY 23 ProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlyLys 42
 DB 347 CCGTTCCTTAGCGCGCGACGCGGAGGAAACAGTAGCGGCAAGCTTCGCGTGTGCTAAG 406
 QY 43 IleMetGluLeuLeuGlyGlnIleGluValGluGluAlaArgGluAlaGlyValIleLeu 62
 DB 407 ATCGTAGCGCTCGGCGGAGAAATGAGTGCATTCGCGAGAAAGAGGTGATCCTCG 466
 QY 63 SerGluAspArgPheTyrIleValIleThrAlaGluGlnLysAlaLysAlaLysGly 82
 DB 467 AGCCAGGATAGCTTCTACCGCTGCTCTTACCTCGGACAGAAAGCCAAAGCCCTGAAGGCG 526
 QY 83 GlnTyrAsnPheAspHisProAspAlaPheAspAsnAspIleuMetHisArgThrLeuLys 102
 DB 527 CAGTTCACCTTGACCCAGCCGATGCTTGAACAATGAACTCATCTCAAAACACTCAAA 586
 QY 103 AsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHisSerArg 122
 DB 587 GAATCACTGAAGGAAACAGTCCAGATCCCGTGTATGACTTGTCTCCACTCCCG 646
 QY 123 LeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIleLeuVal 142
 DB 647 AAGGAGAGACACTTACTGTCTATCCCGACAGCTGTCTCTTTAAGGATCCTGCGC 706
 QY 143 PheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSer 162
 DB 707 TTCTACTCCGAGAGGATGACAGACCTGTCCAGATGAAGCTTTTGTGATACAGATGCG 766
 QY 163 AspValArgLeuSerArgArgValIleLeuArgAspVal---ArgArgGlyArgAspLeuGlu 181
 DB 767 GACACCGGCTCTCAGCAGAGATTAAGGACATGACGACGAGAGCGAGGATCTTGAG 826
 QY 182 GlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeuPro 201
 DB 827 CAGATTTCATCTCAGTACATTAATGCTTCGTCACAGCTGCTTTAGAGAAATTCCTGCGCA 886
 QY 202 ThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsn 221
 DB 887 ACAAGAAGATCTGATGATGATCATCCCTAGAGGTGAGATATCTGGTGGCCATCAAC 946
 QY 222 LeuIleValGlnHisIleGlnAspIleLeuAspGlyAspIleCysLysTrpHisArgGly 241
 DB 947 CTCATCGTCGACGACATCCAGGACATCTGAATGA----- 982
 QY 242 GlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGluLysAspHisProGlyMet 261
 DB 983 -----GGGCGCTCCAAACGCGAGACC-----AATGGCTGTCCAAACGCGCTAC 1024
 QY 262 LeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 DB 1025 ACCCTTCAAGCAAGACGACGATCGAGTCCAGACGAGCGCGCAT 1072
 RESULT 11
 ID ABS51801 standard; cDNA; 734 BP.
 XX AC ABS51801;
 XX DT 21-OCT-2002 (first entry)
 XX DE Human mdtl cDNA Incyte ID No: LI:235557.12:2001JAN12.
 XX KW Human; molecule for disease detection and treatment; MDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; aslma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cyrostatic; antiretroviral; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiaesthetic; gene therapy; gene; ss.

OS Homo sapiens.
 XX
 PN W0200255738-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002MO-US01008.
 XX
 PR 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262326P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 PI Panzer SR, Lincoln SR, Altus CM, Dufour GE, Hillman JL, Jones AL,
 PI Dam TC, Liu TP, Harris B, Flores V, Dafo A, Marwaha R, Chen AJ,
 PI Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
 XX
 DR WPI; 2002-590679/63.
 DR P-PSDB; ABG70328.
 XX
 PT New disease detection and treatment molecule (MDPT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDR expression, e.g. autoimmune or
 PT inflammatory disorders -
 XX
 PS Claim 1; Page 98; 129pp; English.
 XX
 CC The present invention relates to the isolation of novel human
 CC molecules for disease detection and treatment (MDPT), and the
 CC polynucleotide sequences (mdpt) encoding them. The MDDR polypeptides
 CC may be used to screen for molecules that bind to, or are bound by the
 CC encoded polypeptides, and to develop a transcrip image of a tissue or
 CC cell type. Probes comprising at least 20 nucleotides of the mdpt
 CC polynucleotide may be used to assess the toxicity of a test compound.
 CC The MDDR polypeptides and mdpt polynucleotides are useful in the
 CC diagnosis, study, prevention and treatment of diseases associated with
 CC the expression of molecules for disease detection and treatment. Such
 CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
 CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
 CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
 CC polynucleotides may also be used as molecule markers, in microarrays,
 CC and in somatic or germline gene therapy. ABS51779-ABS51814 encode
 CC the MDDR proteins of the invention.
 CC
 XX
 SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
 Alignment Scores:
 Pred. No.: 1,65e-94 Length: 734
 Score: 924.00 Matches: 183
 Percent Similarity: 87.98% Conservative: 0
 Best Local Similarity: 87.98% Mismatches: 4
 Query Match: 63.72% Indels: 22
 DB: 24 Gaps: 1
 US-09-896-522-2 (1-277) x ABS51801 (1-734)
 QY 90 AspAlaPheAspAsnAspLeuMetHisArgThrLeuLeuAsnIleValGluGlyIleThr 109
 DB 14 GATGCTTTGATATATATTGATGACAGACTCTGAAAGACATCGTGGAGGCGCAAAACG 73
 QY 110 ValGluValProThrTyraAspPheValThrHisSerArgLeuProGluThrThrValVal 129
 DB 74 GTGGAGGTGCGGACCTATGATTTTGACACACTCAAGGTTTACAGAGACAGGAGGTC 133
 QY 130 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGluGluIleArg 149
 DB 134 TACCTGCGGAGCGTGGTTCTGTTTGAGGGGATCTTGTTGTCTTACAGCCAGGATCCCG 193

QY 150 AspMetPheHisLeuArgIleuPheValAspThrAspSerAspValArgIleuSerArgArg 169
 DB 194 GACATGTTCCACCTGGGCTCTTCGTGACACCGACTCCGACGTCAAGGCTGTCCAGAG 253
 QY 170 ValLeuArgAspValArgArgIleuArgAspLeuGluGlnIleLeuThrGlnTyrThrThr 189
 DB 254 GTTCTCCGGGAGCGTGGCGCGAGGAGGACCTGGAGCAATTTCTGACGCAATACACAC 313
 QY 190 PheValIleProAlaPheGluGluPheCysLeuProThr----- 202
 DB 314 TTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTGCCGCA-GCAGAGCATCTGACAGGGAAT 372
 QY 203 -----LysIleTyrTrpAlaAspVal 208
 DB 373 GAGAGTCAGCATTTGAGCCCAATGAGTGGTGTGATGAGGGAACAAAGAAATGTCGATG 432
 QY 209 IleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGln 228
 DB 433 ATCATCCCAAGAGAGTGGACATATGTTGCCATCACTGATCGTGCAGCAGCATCCAG 492
 QY 229 AspIleLeuAsnGlyAspIleCysIleTyrPheHisArgGlyGlySerAsnGlyArgSerTyr 248
 DB 493 GACATCTGAATGAGTGCATCTGCAAAATGAGCACCAGAGAGGCTCCCAATGGCGGAGCTAC 552
 QY 249 LysArgThrPheSerGluProGlyAspHisSerGlyMetLeuThrSerGlyLysArgSer 268
 DB 553 AAGCGGACNTTTTCTGAGCGAGGAGCCACCTGGGATGCTGACNTCTGCAAAACGTTAA 612
 QY 269 HisLeuGluSerSerSerArgPro 276
 DB 613 CATTTGAGNCCAGCNCCTGCG 636
 RESULT 12
 ABQ54470
 ID ABQ54470 standard; cDNA; 1322 BP.
 XX
 AC ABQ54470;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HOELP29 cDNA, SEQ ID NO:350.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW anti-inflammatory; gynaecological; reproductive; chromosome 1p32;
 KW gene; ss.
 XX
 OS Homo sapiens.
 PN W0200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001MO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Bires CE, Rosen CA;
 PI WPI; 2002-147878/19.
 DR P-PSDB; ABP41393.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides.

PT disorders such as prostate cancer -
 XX
 PS Claim 1, Page 752, 2338bp, English.
 XX
 CC AAP15566 to AAP16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAP15506 to AAP16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 XX
 XX Sequence 1310 BP; 338 A; 350 C; 352 G; 264 T; 6 other;
 SQ
 Alignment Scores:
 Pred. No.: 9.87e-92 Length: 1310
 Score: 902.50 Matches: 189
 Percent Similarity: 77.82% Conservative: 32
 Best Local Similarity: 66.55% Mismatches: 40
 Query Match: 62.24% Indels: 24
 DB: 21 Gaps: 5
 US-09-896-522-2 (1-277) x AAP15785 (1-1310)
 QY 3 SerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgPro----- 19
 DB 205 GCGAGCGGAGGAGGGGCGCGCGAATGCGCGGAGCAGCAGACCTGCGAGAAC 264
 QY 20 HisGlnArg-----ProPheLeuIleGlyValSerGlyGlyThrAlaSer 34
 DB 265 CACCAGAGCCCAACGGCGCGGAGCCCTTCTTATAGCGCTCAC-GGGGAAACAGCTAGC 323
 QY 35 GlyLysSerThrValCysGlyLysIleMetGluLeuLeuGlyGlnGluValGluGln 54
 DB 324 GCGAAGCTTCCGTCGTGTCATGATCGTACAGCTCTCGGCGAGATGAGTGAGCTAT 383
 DB 55 ArgGlnArgLysValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlu 74
 DB 384 GCGCAGAGAGGAGGATGATCTGAGCGAGATAGCTTACCGTGTCTTACCTCGAG 443
 QY 75 GlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsn 94
 DB 444 CAGAAGGCCAAACCCCTGAAGGSCCACTTCACTTGCACACCCGATGCTTGACAAT 503
 QY 95 AspLeuMetHisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThr 114
 DB 504 GARSTCATCTCTCAACACTCAAGAAATCATGAAAGGGAAGGAAACAGTCCAGATCCCGTG 563
 QY 115 TyrAspPheValIleHisSerArgLeuProGluThrThrValValTyrProAlaAspVal 134
 DB 564 TATGACTTGTGTCTCCATCTCCGAGAGAGAGACAGTTACTGTATCTCCGAGACGTG 623
 QY 135 ValLeuPheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeu 154
 DB 624 GTGCTCTTTGAAGGATCCCTGGCTTCTACTCCAGAGAGTACAGACCTGTTCCAGATG 683
 QY 155 ArgLeuPheValAspThrAspSerAspValArgLeuSerArgArgValLeuArgAspVal 174
 DB 684 AACCTTTTGTGATACAGATGCGGACACCGGCTCTCAGCAGAGATATTAAGGACATC 743
 QY 175 ---ArgArgGlyArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysPro 193
 DB 744 AGCGAAGAGGAGGAGGATCTTGAGCAGATTTTATCTCAGTACATTCAGTTCGTCAGGCT 803
 QY 194 AlaPheGluGluPheCysLeuProThrLysLysThrAlaAspValIleIleProArgGly 213

DB 804 GCCTTGAAGAAATTCTGCTTGCACCAAAAGATATGTGATGATCATCCCTAGAGT 863
 QY 214 ValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGly 233
 DB 864 GCAGATAACTGTGTGCGCATCAACCTCATGTCAGACATCCAGACATCTCGAATGGA 923
 QY 234 AspIleCysLysTyrPheHisArgGlyLysSerAsnGlyArgSerTyrLysArgThrPheSer 253
 DB 924 -----GGGCTCTCAAAACGGCAGACC----- 944
 QY 254 GluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSer 273
 DB 945 ---AATGGCTGTCTCAACCGCTACACCTTACGCAAGAGGAGGAGCATCGAGTCCAGC 1001
 QY 274 SerArgProHis 277
 DB 1002 AGCAGGCCGCAT 1013
 RESULT 14
 AA159636/C
 ID AA159636 standard; cDNA, 1316 BP.
 XX
 AC AA159636;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 3625.
 XX
 KW Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSB-) HYSBQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR MPI; 2001-442253/47.
 DR P-PSDB; AAM40480.
 DR
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1, SEQ ID NO 3625; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as


```
Qy 43 IImecGlureuengIyGInaenGIuValGIuGInuArgIuValIleLeu 62
Db 178 ATCATGAGCAACTTGGGCGAGGATACACGCGCGAGGTGTCCATC 237
Qy 63 SerGInaSpArgPheTyrLysValLeuThrAlaGIuGInuLysAlaLeuLysGIy 82
Db 238 AGTCAGGACAGCTTCACCGGAACTGACGCCCGAGAAAGGCCAGAAAGGC 297
Qy 83 GInTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLys 102
Db 298 CTCTTCAACTTCGATCATCCGATGCCCTTCACGAGAGCTCATGTACAGCAGCTTCAG 357
Qy 103 AsnIleValGIuGIuLysThrValGIuValProThrTyrAspPheValThrHisSer--- 121
Db 358 AACATCTTAAGGCGCCAAAGTGAGATACCCAGCTATGACTACCGACCAACTCTCTC 417
Qy 122 ArgLeuProGIuThrThrValValTyrProAlaAspValValLeuPheGIuGIuLysIleLeu 141
Db 418 GACTTGAAGAACGTTCTGTATCTACCGCCGACCTGCTTGTTCGAGGCGATCTTG 477
Qy 142 ValPheTyrSerGIuGInuIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 161
Db 478 GTATTTTACTTCCCAAGATACGAGAACTGTCACATGAAGCTGTTGTTGACACAGAC 537
Qy 162 SerAspValArgLeuSerArgArgValLeuArgAspVal---ArgArgGIuArgAspLeu 180
Db 538 TCCGATACCGAGGTGGCCAGAGAGTGCACGATGATCAATGACGTTGTCGGGACTTG 597
Qy 181 GIuGInuIleLeuThrGInTyrThrThrPheValLysProAlaPheGIuGInuPheCysLeu 200
Db 598 GATGCGGTGCTCACCCGATATATGACCTTGTAAAGCCGCTTCGAGAGTTCTGCTCG 657
Qy 201 ProThrLysLysTyrAlaAspValIleIleProArgGIuValAspAsnMetValAlaIle 220
Db 658 CCCACGAAAAAGTTGCTGACGTTATTATACACGAGCGCCGCAACACAGTTGCCATT 717
Qy 221 AsnLeuIleValGIuHisIleGIuAspIleLeu 231
Db 718 GATCTTATTTGTACCACTATCGGGAGATTCTC 750
```

Search completed: November 25, 2003, 08:17:56
Job time : 330 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:10:26 ; Search time 3972 Seconds

(without alignments)
2852.962 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450
Sequence: 1 MASAGEDESPAPFADRP.....HPGMLTSGKSHLESSRPH 277

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/runac_21112003_184353_5003/app_query.faeta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blisum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=gat -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896522.@CGN_1.1.3508.@runac_21112003_184353_5003 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_stb: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pac: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_stb: *
28: em_un: *

29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pin: *
35: em_hcg_rod: *
36: em_hcg_mam: *
37: em_hcg_vit: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	834	6 AX449219	AX449219 Sequence
2	1450	100.0	834	9 AF254133	AF254133 Homo sapi
3	1450	100.0	1022	9 AF237290	AF237290 Homo sapi
4	1450	100.0	1624	6 AX449217	AX449217 Sequence
5	1450	100.0	2160	6 BD157613	BD157613 Primer to
6	1450	100.0	2160	9 AK022317	AK022317 Homo sapi
7	1387.5	95.7	2228	9 AK057848	AK057848 Homo sapi
8	1372.5	94.7	1395	9 AF125106	AF125106 Homo sapi
9	1325	91.4	1959	10 BC025146	BC025146 Mus muscu
10	1266	87.3	1810	10 MUSURKI	L31783 Mus musculu
11	1239	85.4	2072	9 BC015547	BC015547 Homo sapi
12	1122	83.6	254933	2 AC078885	AC078885 Mus muscu
13	1159.5	80.0	952	10 MUSURKIA	L31784 Mus musculu
14	1076	74.2	192336	2 AC118474	AC118474 Mus muscu
15	1076	74.2	211075	10 AC076974	AC076974 Mus muscu
16	1074.5	74.1	753	6 BD146824	BD146824 Primer to
17	941.5	64.9	1209	9 BC002906	BC002906 Homo sapi
18	941.5	64.9	1322	6 AX135546	AX135546 Sequence
19	934	64.4	1121	9 AF236637	AF236637 Homo sapi
20	934	64.4	1312	10 BC023789	BC023789 Mus muscu
21	934	64.4	1354	10 AF236636	AF236636 Mus muscu
22	932.5	64.3	744	9 AB062451	AB062451 Homo sapi
23	924	63.7	734	6 AX540411	AX540411 Sequence
24	909.5	62.7	1784	5 BC045968	BC045968 Danio rer
25	733	50.6	1467	3 AY119583	AY119583 Drosophil
26	656.5	45.3	147492	9 AL358781	AL358781 Human DNA
27	656.5	45.3	241882	10 AL808027	AL808027 Mouse DNA
28	656.5	45.3	254677	2 AC068494	AC068494 Mus muscu
29	631.5	43.6	244105	2 AC098897	AC098897 Rattus no
30	622.5	42.9	156905	2 AC020006	AC020006 Drosophil
31	622.5	42.9	199016	3 AC008204	AC008204 Drosophil
32	622.5	42.9	206741	3 AB003747	AB003747 Drosophil
33	572.5	39.5	136240	3 AC117070	AC117070 Dictyoste
34	570.5	39.3	1492	8 BT002336	BT002336 Arabidops
35	570.5	39.3	1844	8 AY089970	AY089970 Arabidops
36	545.5	37.6	1847	6 AX463241	AX463241 Sequence
37	545.5	37.6	1847	9 AK000524	AK000524 Homo sapi
38	544	37.5	1401	6 AX507248	AX507248 Sequence
39	544	37.5	1432	8 AY122946	AY122946 Arabidops
40	544	37.5	1785	8 AY072218	AY072218 Arabidops
41	542	37.4	3246	8 AY129436	AY129436 Drosophil
42	540.5	37.3	1813	6 AX405842	AX405842 Sequence
43	540.5	37.3	1875	6 BC033078	BC033078 Homo sapi
44	539.5	37.2	1823	10 BC016535	BC016535 Mus muscu
45	529	36.5	1831	8 AY086133	AY086133 Arabidops

RESULT 1

ALIGNMENTS

AX449219
 LOCUS AX449219 834 bp DNA linear PAT 03-JUL-2002
 DEFINITION Sequence 3 from Patent WO0202761.
 ACCSSION AX449219
 VERSION AX449219.1 GI:21697996
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Gluckemann, M.A.
 57658, a human uridine kinase and uses thereof
 JOURNAL Patent: WO 0202761-A 3 10-JAN-2002;
 Millennium Pharmaceuticals, Inc. (US)
 Location/Qualifiers

FEATURES
 source 1..834
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 194 a 218 c 265 g 157 t

ORIGIN

Alignment Scores:
 Pred. No.: 4.11e-129 Length: 834
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-896-522-2 (1-277) x AX449219 (1-834)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 1 ATGGCTTCGGCGGAGCGAAGACTCGAGAGCCCGCGCGAGCGACCGTCCGAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40
 DB 61 CAGCGCCCTTCCTGATAGGGGTAGCGCGCGACTCGACGGGAAAGTCGACCGTGT 120

QY 41 GlnValIleMetGluLeuGlyValAsnGluValGluGlnArgGlnArgValVal 60
 DB 121 GAGAGATCATGAGTCTGCTCGGACAGACAGAGTGAACGCCGACGGAAGTGTCTC 180

QY 61 IleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGluGlnValAlaValLeu 80
 DB 181 ATCTGAGCCGACGACAGGTTCTCAAGGTCTCTGACGACAGACGAGGACCTTG 240

QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 241 AAAGAGACAGTAAATTTTGAACATCCAGATCCCTTTTAATGATTTGATGACAGGACT 300

QY 101 LeuValAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 301 CTGAAGAACATCTGCGAGGCGCAAAACGCTGAGAGTCCGACCTTATTTTGACACAC 360

QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 DB 361 TCAAGGTACACAGACACACGCTGCTACCTCGGACGAGTGTCTGTTGAGGCGCATC 420

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
 DB 421 TTGGTGTTCACACGACGAGATCCGGACATGTTCCACCTGCGCTCTTCTGTGACACC 480

QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 DB 481 GACTCCACAGCTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCGACGAGGAGGACTTG 540

QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 541 GAGGAGATTTCGACGACGATACACACCTTGTGAAGCGGCTTCGAGGAGTTCTGCTG 600

QY 201 ProThrLysLysThrValAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 DB 601 CCACCAAGAGATATCCAGATGATCATCCACGAGAGTGAACATATGTTGCCATC 660

QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 DB 661 AACCTAGTGTGACGACATCCAGACGATCTCGAATGATGATCTCGAATGACACCGA 720

QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
 DB 721 GGAGGTCACAAATGGCGGAGCTCAACGCGACCTTTCTGAGCCAGGAGCACACCTTG 780

QY 261 MetLeuThrSerGlyLysArgSerHisIleuGluSerSerArgProHis 277
 DB 781 ATCTGACCTCTGCGAAGCGTCACTTTGAGTCCAGACGACCCGAC 831

RESULT 2
 AF254133 834 bp mRNA linear PRI 02-MAY-2001
 LOCUS AF254133
 DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
 ACCSSION AF254133
 VERSION AF254133.1 GI:13924749
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Human uridine kinase from prostate cancer cell line (LNCap)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 834)
 Ho, Y.S. and Johnson, R.K.
 Direct Submision
 Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709
 Swedeland Road, King of Prussia, PA 19406, USA
 Location/Qualifiers

FEATURES
 source 1..834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..834
 /note="protein kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AAK49122.1"
 /db_xref="GI:13924750"
 /translation="MASAGDEPCESPAPDRPHQRPLIGVSGTASGKSTVCEKIM
 ELIGONEVEORORKVYILSDREFKVLTEORAKALKGOYNPDPAFPNDLMHRTLK
 NIVEGTVVADPVTHSRLEPTTVYVADVLVPEGLIVFSOEIRDMFHLRPVDT
 DSDVRLSRVLDVDRGRDLEQLITQYTFYVAPAREPCLPTKKYADVITIPGVDMV
 AINLVQHIQDILINDICMKHNGSGNSGRYKRTFSPGHPMLISGKSHLESSRP
 H"

BASE COUNT 194 a 219 c 265 g 156 t

ORIGIN

Alignment Scores:
 Pred. No.: 4.11e-129 Length: 834
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-896-522-2 (1-277) x AF254133 (1-834)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 1 ATGGCTTCGGCGGAGCGAAGACTCGAGAGCCCGCGCGAGCGACCGTCCGAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40
 DB 61 CAGCGCCCTTCCTGATAGGGGTAGCGCGCGACTCGACGGGAAAGTCGACCGTGT 120

[illegible]

FEATURES	source
AUTHORS	Van Rompay, A. R., Linden, K., Norda, A., Zhu, C., Zheng, X., Johansson, M. and Karlsson, A.
TITLE	Direct Submission
JOURNAL	Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge University Hospital, Stockholm 14186, Sweden
location/Qualifiers	
gene	1..1022 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"
CDS	1..1022 /gene="VCK1" 5..838 /gene="VCK1" /function="phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate" /codon_start=1 /product="uridine-cytidine kinase 1" /protein_id="AAK28324.1" /db_xref="GI:13506765" /translation="MASAGEDCSPEADRPDRPDRPFLIGSGGTASGSTVCEKIM ELGCGNEORQRKVLLSDRPYKULTAQOKALGVGNFHPDAFNDIMRLTK NIVEKTVFPTDFTVTHSRLEPFTTVYPRDVLFEELVYSEIDIMHRLPDTT DSVRLSRVLRVDRGRDELQILQTYTTTVPKAFEEFLPTKYLAVVLIIPGVDMV AINIIVIOHIDILNGDICKMHRGSGNSYKRTFSEPDHGMILTSKRSHLESSRR H"
BASE COUNT	227 a 282 c 324 g 189 t
ORIGIN	
Alignment Scores:	
Pred. No.:	5.33e-129 Length: 1022
Score:	1450.00 Matches: 277
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	9 Gaps: 0
US-09-896-522-2 (1-277) x AF237290 (1-1022)	
QY	1 MetAlaseeraIaGlyGlyGluAspCYeGluSerProAlaProGluAlaAspArgProHis 20
Db	5 ATGGCTTCGGCGGGAAGCGAAGACTGCGAGAGCCCCCGCGGAGCGCGACCGTCCGAC 64
QY	21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db	65 CAGCGGCCCTTCCTCGATAGGGGTGAGCGCGCGGCGACTCCAGCGGAGATCGACCGTGTGT 124
QY	41 GluIuylleMetGluIuLeuLeuGlyGluAsnGluValIuGluIuAsnGluIuAsnGlyValVal 60
Db	125 GAGAGAGTCATGAGGAGTGTCTGGGCAACAACAGGTGTGAACAAGCGGACGGAAAGTGTCTC 184
QY	61 IleuSerGlnAspArgPheThrValValleuThrAlaGluGluIuValAlaValAlaIleu 80
Db	185 ATCTGAGCCAGCAGACAGTCTTAACAAGTCTCTGACGCGCAGCAGCAAGAGGCCAAGCCTTG 244
QY	81 LysGlyGluIuThrAsnPheAspHisPProAspAlaPheAspAsnAspLeuMetHisAlaGThr 100
Db	245 AAAGGACAGTACAAATTTTGAACATCCAGATGCCCTTTATATGATGTTTGAAGTCACAGACT 304
QY	101 LeuIuAsnIleValIuGlyLysThrValGluValProThrTYrAspPheValThrHis 120
Db	305 CTGAAGAACATCGTGAAGGGCAACAGGTGAGGTGCCAGCTATGATTTGTGACACAC 364
QY	121 SerArgLeuProGluIuThrValValIuTYrProAlaAspValValleuPheGluGlyIle 140
Db	365 TCAAGGTTTACAGAGACCAAGGTGTCTTACCTTCGAGCAGCGGTCTGTTTGAAGGCATC 424
QY	141 LeuValPheTYrSerGluGluIleArgAspMetPheHisleuArgLeuPheValAspThr 160
Db	425 TTGGATGTTTACAGCCAGAGATCCGGGACATGTTTCACTCGGCGCTTTCGTGTGAACAC 484
QY	161 AspSerAspValArgLeuSerArgArgValleuArgAspValAlaArgArgIuArgAspLeu 180

DB 485 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCCGAGGAGGAGACTG 544
 QY 181 GUGUlnlleuThrGlnTyrThrPheValIysProAlaPheGluGluPheCysLeu 200
 DB 545 GAGCAGATTCTGACGCGACTACACACCTTCCTGGAAGCCGCTTCGAGGAGTTCTGCTG 604
 QY 201 ProThrIysIysTyrAlaAspValIleIleProArgIValAspAsnMetValAlaIle 220
 DB 605 CCGCAAGAAGATATGCCGATGTATCATCCCGGAGAGTGAACAATATGTTGCAATC 664
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTyrHisArg 240
 DB 665 AACCTGATCGTGACGACATCCAGACATCTGATGATGTCATCTGCAATATGCGACCGA 724
 QY 241 GlyGlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAspHisProGly 260
 DB 725 GAGAGGCTCCATGCGGAGCTGACAGAGGAGCACTTTCTGAGCCAGGAGGACCACTTG 784
 QY 261 MetLeuThrSerGlyIysArgSerHisIleuGluSerSerArgProHis 277
 DB 785 ATGCTGACCTCTGGCAACGGTCAATTTGGAGTCCAGACAGACCCAC 835
 RESULT 4
 AX449217 1624 bp DNA linear PAT 03-JUL-2002
 LOCUS AX449217
 DEFINITION Sequence 1 from Patent WO0202761.
 ACCESSION AX449217
 VERSION AX449217.1 GI:21697994
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Glucksmann, M.A.
 TITLE Patent: WO 0202761-A 1 10-JAN-2002;
 JOURNAL Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 SOURCE Location/Qualifiers
 1..1624
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 94..937
 /note="unamed protein product"
 /codon_start=1
 /protein_id="CAD37771.1"
 /db_xref="GI:21697995"
 /translation="MASAGBDCSPAPADRPHPORPFLIGVSGTASGKSTVCEKIM
 ELIGONEVEORQRKVVILSDRFYKVLTAEGKAKALGOVNFDPDAFNDIMHRTLK
 NIVGKTEVEPTTYPVTHSRLEPTTVYPPADVYLFESILVPSSEITDMFLHLPVDT
 DSDVLSRRVLRKDRGDLQILITQITTFKAPFESECLPTKTAIVYIIPRGDNNV
 AINLVHIDILINDGICIKMHRGSSNRSYKRPSEPDHGMLTGSKRSHLESSRP
 H"
 CDS
 BASE COUNT 354 a 427 c 506 g 337 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.65e-129 Length: 1624
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-896-522-2 (1-277) x AX449217 (1-1624)
 QY 1 MetAlaSerAlaGlyGlyAspCysGluSerProAlaProGlyAlaAspArgProHis 20
 DB 94 ATGGCTTCGGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 153
 QY 21 GlnArgProPheLeuIleGlyValSerGlyIlyThrAlaSerGlyIysSerThrValCys 40

DB 154 CAGGGGCCCTTCCGTATAGGGGTGAGCGGGGACCTGCCAGCGGAGATCGACCGTGTGT 213
 QY 41 GlnIlyIleMetGluLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgIysValVal 60
 DB 214 GAGAAATCATGATGATGTTGCTGGGACGAAAGAGTGTGAACAGCGGAGCGGAGGTTG 273
 QY 61 IleLeuSerGlnAspArgPheTyrIlyValIleThrAlaGlnIlyValIleValIle 80
 DB 274 ATCTTAGCCGAGACAGGTTCTTAAGAGTCTGACGGCAGAGCAAGGCCAAGGCTTG 333
 QY 81 IysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 334 AAGGACAGTACATTTTGACCATCCAGATGCTTGATTAATGATTTGATGACAGAGCAT 393
 QY 101 LeuIlyAsnIleValGluGlyIysThrValGluValProThrTyrArgPheValThrHis 120
 DB 394 CTGAAGAACATCTGTGAGGCGCAAAAGGAGAGTGTCCACCTATATTTTGTGACACAC 453
 QY 121 SerArgLeuProGluThrThrValIleTyrProAlaAspValIleuPheGluGlyIle 140
 DB 454 TCAAGTTACCAAGACACCGGTGTGTCTACCTCGCGAGCTGTTCTGTGAGGCGATC 513
 QY 141 IleuValPheTyrSerGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
 DB 514 TTGGTGTCTTACAGCGCAGGAGATCCGGACATGTTCCACTGGCGCTTCGTGAGACACC 573
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 DB 574 GACTCCGACGTCGGGTGTCTCGAAGAGTTCTCCGAGAGTGGCCGAGGAGGAGCACTG 633
 QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
 DB 634 GAGCAGATTCTGACGACGATACACCACTTCGTGAACCCGCTTCAGAGAGTTCTGCTC 693
 QY 201 ProThrIysIysTyrAlaAspValIleIleProArgIValAspAsnMetValAlaIle 220
 DB 694 CCGACAAAGAAGATGCCGATGTATCATCCACGAGAGTGTGAACAATATGTTGCCATC 753
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIysTyrHisArg 240
 DB 754 AACCTGATCGTGACGACATCCAGACATTCGAAATGTGACATTCGCAATGTGCACCGA 813
 QY 241 GlyGlySerAsnGlyArgSerTyrIlyArgThrPheSerGluProGlyAspHisProGly 260
 DB 814 GAGGCTCCATGCGGAGCTGACAGCGGACCTTTCTGAGCCAGGAGACCACTCTGG 873
 QY 261 MetLeuThrSerGlyIysArgSerHisIleuGluSerSerArgProHis 277
 DB 874 ATCTGACCTCTGGCAACGGTCAATTTGGAGTCCAGACAGACCCAC 924
 RESULT 5
 BD157613 2160 bp DNA linear PAT 17-JAN-2003
 LOCUS BD157613
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BD157613.1 GI:27863371
 VERSION JP 2002191363-A/12456.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2160)
 TITLE Primer for synthesizing full-length cDNA and use thereof
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ienli,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
 Patent: JP 2002191363-A 12456 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/12456
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KEIICHI NAGAI, TETSUJI OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/00, C12N5/00, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof PH Key
 Location/Qualifiers
 (95) : (925) .
 FT CDS Location/Qualifiers
 1. 2160
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 457 a 591 c 671 g 441 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.39e-128 Length: 2160
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-896-522-2 (1-277) x BD157613 (1-2160)
 QY 1 MetLAserrAagLygluAapPcygluSerProAlaProgluAlaAapAapProHis 20
 Db 95 ATGGCTTCGGGGGAGGCGAAGATCGGAAGCCCCCGCGGCGCGCCGCTCCAC 154
 QY 21 GlAapPpheLeuilegluValSerGlYglYThraLaserGlYvsaerThraValCys 40
 Db 155 CAGCGGCCCTTCCTGATAGGGGCTGAGCGCGGCACTCCGACGGGGAAGTGCAGCGTGT 214
 QY 41 GluLysileMetgluLeuLeuGlYgluValAsnGlYgluValLeuArglnArglnValVal 60
 Db 215 GAGAGATCATGAGGCTGCTGGGACAGACAGAGTGAACAGCGGCGCGGGAAGGTGTC 274
 QY 61 lIleuSerGlAspAapPheTyrlYsValleuThraLagluGlnlysaAlaValleu 80
 Db 275 ATCTTGCCAGACAGAGCTTCTACAGCTCTGACGCGCAGAGCGCAAGCGCCCTTG 334
 QY 81 LySGlYgluTyraSpPheAapHisProAapAlaPheAapAapPheMetHisArgThr 100
 Db 335 AAGAGACAGTACATTTTGCATCCAGATGCCCTTGATATGATTTGATGCACAGACT 394
 QY 101 leuLysAsnileValgluGlYlYsThraValgluValProThraSpPheValThHis 120
 Db 395 CTGAGAGACATCGGAGGCGCAAAACGCTGAGGTGCCGACCTATGATTTTGTGACACAC 454
 QY 121 SerArgleuProgluThraValValTyrlProAlaAapValleuPhegluGlYlYle 140
 Db 455 TCAAGGTTCACAGACACAGCGTGTCTACCTCGCGGACGCTGTTGTTGAGGGGATC 514
 QY 141 leuValPheTyrlSerGlYgluValleuArgAapPheHisleuArgleuPheValAapThr 160
 Db 515 TTGGTGTTCACAGCGGAGGATCCGAGACATGTTCCACTGCGCTCTTGTGAGACACC 574
 QY 161 AspSerAapValArgleuSerArgValleuArgAapValArgArglYAspAapLeu 180
 Db 575 GACTCCGAGCTCAGGCTGCTCGAAGATCTCCGAGGACGTGCGGAGGAGGAGACTG 634
 QY 181 GluGlnleuThraGlnTyrlThraPheValleuProAlaPhegluGluPheCysLeu 200
 Db 635 GAGCAGATTCGACGAGTACACCACTTGTGAAGCGCGCTTCGAGGAGTTCGCTG 694
 QY 201 ProThraLysTyrlAaAapValleuProArgGlYValAaAapAapValleu 220
 Db 695 CCGACAAAGAGATGCGGATGATCATCCACGAGGAGTGAACATATGATGCGCATC 754
 QY 221 AsnleuileValGlnHisileGlnAapileleuAapGlYAspilleCyslystPHisArg 240

Db 755 AACCTGATCGTCAGACATCCAGACATTCGATGATGATCATCTGCAATGAGCACCA 814
 QY 241 GlYglYserAsnGlYArgserTyrlYsArgThraPheSerGlulProglYAspHisProglY 260
 Db 815 GGAGGGTCCAAATGGGGGAGCTTCAAGCGGACCTTTCTGAGCGAGGAGCACCTGGG 874
 QY 261 MetLeuThraSerGlYvsaAapPheHisleuGlYvsaAapPheHisArgProHis 277
 Db 875 ATCTGATCTCTGCGCAACCGTCACTTTGAGTCCAGACGAGACCCAC 925
 RESULT 6
 AK022317 2160 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ12255 fis clone MAMMA1001476, highly similar
 DEFINITION to URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK022317 GI:10433687
 VERSION AK022317.1
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Bukaryoca, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1
 Isogai, T., Ota, T., Hayaishi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
 Nishikawa, T., Nagai, K., Sugano, S., Shitatori, A., Sudo, H.,
 Magatsuna, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, W., Chiba, Y., Ishida, S., Murakawa, K.,
 Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
 Ishii, S., Kawai, Y., Saito, K., Yamamoto, Y., Wakamatsu, A.,
 Nakamura, Y., Nagahara, K., Masuo, Y., Ninomiya, K. and Iwayanagi, T.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 2160)
 REFERENCE Isogai, T. and Otsuki, T.
 AUTHORS Direct Submission
 TITLE Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
 JOURNAL Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 source
 1. 2160
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MAMMA1001476"
 /issue_type="Mammalian"
 /clone_id="MAMMA1"
 /note="cloning vector: pME18SFL3"
 95. 928
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB14010.1"
 /db_xref="GI:10433688"
 /translation="MASAGGEDCESAPAPADPHORPFLIGVSGGASGKSTVCEKIM
 ELIGNEVEORRKYVILISODRFYVLTAEOAKAKLKGQYNDHDAFNDPMHRTIK
 NIVEKTYEVPYPTVTHSRLEPETYVYPADVLEGLIVFESQIRDMFHLRLVDT
 DSDFVRSRLVDVRRGLDQILTYITFVAPAEFELPYTKTADVTIIPGVDMV
 AINLVQHTQDILNDGICKWHKGSNGRSYKRTFSEPDHPMLTSGKRSIHSSSRP
 H"
 BASE COUNT 457 a 591 c 671 g 441 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.39e-128 Length: 2160
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-896-522-2 (1-277) x AK022317 (1-260)			

OY	1	MetAlaSerAlaGlyGlyGluIuAspCysGluSerProAlaProGluIuAlaAspArgProHis	20
Db	95	ATGGCTTGCGCGGAGGCGAAGCTCGCGAGGCCCCGCGCGGAGGCGCGACTCCGAC	154
OY	21	GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys	40
Db	155	CAGCGGCCCTTCTCTGTAGTGGGGAAGCGGCGGCGACTGCCAGCGGGAAGTCGACTGTGT	214
OY	41	GlyLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal	60
Db	215	GAGAAAGATCATGAGACTGCTGGGACAGAAACGAGTGGAAACGCGGACGCGAAGGTGGTC	274
OY	61	IleLeuSerGlnAspArgPheThrLysValIleThrAlaGluGlnLysAlaLysAlaLeu	80
Db	215	ATCTCGAGCCAGAGACAGGTTCTTCAAGGTCCTGACGCGACAGACAAAGCCAGGCTTGG	334
OY	81	LysGlyGlnIlyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
Db	335	AAAGGACAGTACCAATTTGACCATCCAGATGCCCTTGATATGATGATTTGATGCACAGACT	394
OY	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrLysAspPheValThrHis	120
Db	395	CTGAAGAACATGTGGAGGGGCAAAAGGAGGAGGTCCGACCTATATTTTGGACACAC	454
OY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140
Db	455	TCAAGSTTACAGAGAACCCAGGTGGCTTACCTCGCGAGCGTGGTTCTGTTTGAAGGCATC	514
OY	141	LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
Db	515	TTGGTGTTCTTACAGCCAGAGAGATCCCGGACATGTTCCACTCGGCGCTTCTCGTGGACACC	574
OY	161	AspSerAspValArgLeuSerArgValIleuArgAspValAlaArgGlyValArgAspLeu	180
Db	575	GACTCCGACGTAGAGGTGTCCTCGAAGAGTTCTCCGGAACGTGGCGGAGGAGGACCTG	634
OY	181	GluGlnIleLeuThrGlnIlyrThrThrPheValLysProAlaPheGluGluPheCysLeu	200
Db	635	GAGCAGATTCTGACCGAGTACACCACTTCGTGAAGCCGGCTTCGAGGAGTTCTGCTCG	694
OY	201	ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle	220
Db	695	CCGACAAAGAAAGTATAGCCGATGTGAATCCACCGAGGAGTGACAAATATGTGTGCATC	754
OY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTTPHisArg	240
Db	755	AACTCGATCTGCAGACATCCAGAGCATTTCTGAAATGGTACACTTCTGCAAATGGCACCGA	814
OY	241	GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyLysAspHisProGly	260
Db	815	GGAGAGTCCAAATGGGCGAGTACAAAGCGGACCTTTCTGAGCCAGGGGACCACTCGGG	874
OY	261	MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis	277
Db	875	ATGCTGACCTCTGGCAAAAGCTGACATTTGGAGTCCAGAGCAGACCCAC	925

RESULT 7	
AK057848	
LOCUS	2228 bp mRNA linear PRI 31-OCT-200
DEFINITION	Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to URIDINE KINASE (EC 2.7.1.48).
AK057848	
AK057848.1	GI:16553809
VERSION	oligo capping: fis (Full insert sequence).
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS

REFERENCE AUTHORS	TITLE	JOURNAL REFERENCE AUTHORS	JOURNAL TITLE
1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, T., Komiyama, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isegai, T. and Sugano, S.	NEDO human cDNA sequencing project		
	unpublished		
	2 (bases 1 to 2228)		
	Sugano, S. and Suzuki, Y.		
	Direct Submission		
	Submitted (24-OCT-2001)		
	Sumio Sugano, Institute of Medical Science,		

COMMENT

COMMENT
NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo; Laboratory of genome structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES
SOU

```

source
1. .2228
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR05878"
/tissue_type="brain"
/clone_id="CBR"
/notes="Cloning vector: pME18SFL3"
478 a 598 c 696 g 456 t
BASE COUNT
```

Alignment Scores:	
Pred. No.:	1,33e-122
Score:	1387.50
Percent Similarity:	88.75%
Best Local Similarity:	88.75%
Query Match:	95.63%
Gaps:	9
DB:	2
length:	2238
Matches:	276
Conservative:	0
Mismatches:	8
Indels:	35
Gaps:	2

US-09-896-522-2 (1-277) X AK057848 (1-2228)

OY	1	MetAlaSerAlaGlyGlyGluInsProCysGluSerProAlaProGluAlaAspArgProHis	20
Db	64	ATGGCTTCGGCGGAGCGCAGACCTGCAGAGCCCCGGCGGAGCGCACCGTCTGCAC	123
OY	21	GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLys-----	36
Db	124	CAGCGGCCCTTCTCTGATAGGGGTGACGGCGCACTGCACGGGAAAGCACACGGGTTCC	183
OY	37	-----SerThrValCysGluLysIleMetGluLeu	46
Db	184	CAGCGGGAATCCCGCTCTGTCTCCAGTCGACCGGTGTGAGAAAGATCATGAGATTGG	243
OY	47	LeuGlyGlnAsnGluValGluGlnArgGlnArgLysValIleLeuSerGlnAspArg	66
Db	244	CTGGGACAGAACCGAGGTGGAGACGGGAGCGGAGGGTGTCATCTCGAGCGACGAGACGG	303
OY	67	PheTyrLysValLeuThrIleGluGlnLysValLysAlaLeuLysGlyGlnTyrAsnPhe	86
Db	304	TTCTTACAGGTCCTGACCGCAGACGAAAGGCCAAGGCTTGAAAGACACATGACATTTT	363
OY	87	AspHisProAspAlaPheAsnAsnAsnLeuMetHisArgThrIleLysAsnIleValGlu	106
Db	364	GACCATCCAGATGCCCTTGTATGATGATTGTATGCACAGACTTGAAAGAACATGCTGGAG	423
OY	107	GlyLysThrValGluValProThrTyrArgPheValThrHisSerAlaGluLeuProGluThr	126
Db	424	GGCAAAACGGTGGAGGTGGCGACCTATATTTTGGACACCTCAAGGTTAACCGAAGACC	483

QY 127 ThrValValTyrProAlaAspValValIleuPheGluGlyIleValAlaPheTyrSerGln 146
 DB 484 ACGGTGCTACCTCGGAGAGTGTCTGTTTGAAGGACATCTTGTTCTACAGCCAG 543
 QY 147 GluIleArgAspMetPheH1sLeuArgLeuPheValAspThrAspSerAspValArgLeu 166
 DB 544 GAGATCCGGGACATGTTCCACCTCGCTCTCTGTCGACACCGACTCGAGCTGAGCTG 603
 QY 167 SerArgArgValIleuArgAspValArgArgGlyValArgAspLeuGluGlnIleLeuThrGln 186
 DB 604 TCTCGAAGAGTCTCCCGGAGGTGCGCGAGGGAGGACCTCGAGCAGATTCTGACCCAG 663
 QY 187 TyrThrThrPheValIleuPheGluGluGluPheCysLeuPheProThr----- 202
 DB 664 TACACCACTTCGTGGAAGCCGCTTGTGAGAGATTCTGCTGCTGCCA-GAGCATCTGACA 722
 QY 203 -----LysIleTyrAla 206
 DB 723 GGGAAATGAGATGACATGAGCCATGAGTGTGATGAGGAGACAAAGAGTATGCC 782
 QY 207 AspValIleIleProArgGlyValAspAspMetValAlaIleLeuIleValGlnHis 226
 DB 783 GATGTGATCATCCACAGAGAGTGGACAAATATGTTGCCATCAACCTGATCTGACACAC 842
 QY 227 IleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheIleArgGlyIleSerAsnGlyValArg 246
 DB 843 ATCCAGAGACATCTGAAATGTGATCTGCAATATGCAACGAGAGAGGTCCATATGGCCG 902
 QY 247 SerTyrIleArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerGlyLys 266
 DB 903 AGCTACAGACGACCTTTCTGAGCCAGGGGACCACTCGGATGTCGACTCTGCGAAA 962
 QY 267 ArgSerHisLeuGluSerSerSerArgProHis 277
 DB 963 CGGTCACTTTGGAGTCCAGACAGACCCAC 995

RESULT 8
 AF125106 1395 bp mRNA linear PRI 07-FEB-2002
 LOCUS Homo sapiens uridine kinase mRNA, complete cds.
 DEFINITION AF125106
 ACCESSION AF125106
 VERSION AF125106.1 GI:18568108
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1395)
 Xing, Y.R., Yu, L. and Zhao, S.Y.
 Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1395)
 Ding, J.B., Yu, L. and Zhao, S.Y.
 Direct Submission
 TITLE Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433, People's Republic of China
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. 1395
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 8. 838
 /codon_start=1
 /product="uridine kinase"
 /protein_id="AA175943.1"
 /db_xref="GI:18568109"
 /translation="MASAGSEGARARAGANRPHORPPLIGSGTASGKSTVCEKIME
 LIGQNEVQHGKRVILSDRFYKVLAEAKALAKQVFPDPAFDMDLMRILKN
 IVBQKVEPTVPTVTHSRLEPTVVYPADVVEGLVYVSOEIRDMFRLRPVDT
 SDVLRVRVLRGRGDLQILQYTTFFKPAFEBCLEPTKYLAVIIPRGVNNVA
 INLIQHIQDILNDICMKHRRGSGNGRTYKRTSEPDHGMILSGKRSHLESSRPH

BASE COUNT 325 a 352 c 411 g 307 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,97e-121 Length: 1395
 Score: 1372.50 Matches: 270
 Percent Similarity: 98.19% Conservative: 2
 Best Local Similarity: 97.47% Mismatches: 5
 Query Match: 94.66% Indels: 3
 DB: Gaps: 0
 US-09-896-522-2 (1-277) x AF125106 (1-1395)
 QY 1 MetAlaSerAlaGlyGlyAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 8 ATGCTTCCGCGGAGAGCGAAGG-TCGAGAG-CCGCGCGGAGAGG-AACTCTCCGAC 64
 QY 21 GluArgProPheLeuIleGlyValSerGlyIleThrAlaSerGlyLysSerThrValCys 40
 DB 65 CAGCGTCCCTTCTGATAGGGGTAGCGGGGACCTGCCAGCGGAGATGACCGTGTGT 124
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValAla 60
 DB 125 GAGAAGTATGAGATTGCTGGGACAGAACGAGGTGGAACAGCGGACGGTATGGTGTG 184
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 DB 185 ATCTGAGCCAGAGACAGTGTCTCAAGGTCTTGACCGAGAGACAGAACGAGGCTTG 244
 QY 81 LysGlyGluIleTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 245 AARGAGACATGACATTTTGACATCCAGATGCTTGATGATGATTTGATGACAGACT 304
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValIleHis 120
 DB 305 CTGAAGACATGCTGGAGGCAAAACGTTGAGGTGCCACCTATGATTTTGTGACACAC 364
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValAlaLeuPheGluGlyIle 140
 DB 365 TCAGGTTCACAGAGACACAGGTGTCTACCTGCGGAGCGTGTCTGTTGAGGCGCATC 424
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 425 TTGTTTCTTACAGCCAGAGAGATCCGGGACATGTTCCACTGTGCTTCTTGACACAC 484
 QY 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgLysArgAspLeu 180
 DB 485 GATTCGAGCTGACGCTGTCTCGAAGAGTTCTCCGGGAGCTGCGCGGAGGAGACTG 544
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 545 GAGCAGATTCTGACGAGTACACACACTTCGTGAACCGGCTTCGAGAGATTCTCCCTG 604
 QY 201 ProThrLysLysTyrLysAspValIleIleProArgGlyValAspAspMetValAlaIle 220
 DB 605 CCAGCAAAAGAGATGCCATGTGATCATCCACGAGAGTGGACAATATGTTGCCATC 664
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheIleArg 240
 DB 665 AACCTGATGTGAGAGACATCCAGAGCATTTCTGAAATGTGACATCTGCAAAATGGACCGA 724
 QY 241 GlyGlySerAsnGlyValArgSerTyrLysArgThrPheSerGluProGlyLysHisProGly 260
 DB 725 GAGAGGTCCAAATGGGCGAGTACAAAGCGAGACTTTTCTGAGCCAGGGGACCACTCGGG 784
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluLeuSerSerArgProHis 277
 DB 785 ATGTCACTCTGGCAAAACGTTACATTTTGAGTCCAGACGAGACCCAC 835

RESULT 9
 BC025146 1959 bp mRNA linear ROD 16-APR-2003
 LOCUS BC025146

DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
MGC:36231 IMAGE:4913412), complete cds.
ACCESSION BC025146
VERSION BC025146.1 GI:19263563
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1959)
REFERENCE 1 (bases 1 to 1959)
AUTHORS Strauberg, R.L., Feringold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stepleton, M., Soares, M.B., Bonaligo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.D., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gnaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Farey, J., Heiton, E., Kettelman, W., Madan, A., Young, A.C., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.F., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
12477932
2 (bases 1 to 1959)
AUTHORS Strauberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffery E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisgeed, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavali,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1..1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1959
gene

CDs
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AA025146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translations="MAGAGGSGSAAAPADRPQRPPLIGVSGGTASGKSTYCEIM
ELIQNEVBRQRKVLITSDCFYKLTAEQAKAKGQNFDPHAFDNDLHKTLK
NIVEGTSREVPYDFTVTHSRLEPTTVTVADVVLEFGILVFTQELRDMFHLFVDT
DSVDRSLRDLVQGRDLQELITQYTAFAVKAPEEFCLPTKYADVII PRGVDMV
AIVLIQIHDILNGDLCKRHRRGPNRHNKRTFPPBDHPGLATGKRSLESSRP
H"
BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 1,03e-116 Length: 1959
Score: 1325.00 Matches: 254
Percent Similarity: 95.67% Conservative: 12
Best Local Similarity: 91.70% Mismatches: 12
Query Match: 91.38% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x BC025146 (1-1959)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 57 ATGGCTTCGGCGGAGGCGGCGGCTCGAGAGCGCGCGCCGAGCGGATCGTCCCCAG 116
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 117 CCGGGCCGTTCTCATCGCGTGAAGGGGGGAGCCGCTAGTGCAAGTCAACAGTGTGT 176
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 177 GAGAGATCATGAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 236
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 237 ATCTGAGCCAGACATGCTTCTCAAGGTTCTGACCGCGAGAGAGAGAGAGAGAGAG 296
QY 81 LysGluGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 297 AAGGACAGTACATTTTGAACCCAGATGCTTTGATATGATCATGATGACCAAGACC 356
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 357 CTGAAAACATGTTTGAAGCAAAACGTGCGAGAGTCCCTACCTATGATTTTGTGACCCAC 416
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 417 TCAAGGTTTACCAAGACACATGTTGTTACCCAGCTGATGTGTTGTTGAGGGGACATC 476
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 477 TTGTATTCTACACCCAGAGATCCGGAGCATTTTCACTGCGCTCTTTTGGGACACA 536
QY 477 TTTGTATTCTACACCCAGAGATCCGGAGCATTTTCACTGCGCTCTTTTGGGACACA 536
QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 537 GACTCGATGTATGAGGTGTCTGCAAGAGTTCTCGGGAGTGTGCAAGAGAGAGAGAGAG 596
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 597 GAGCAGATCTGACCTGACAGTACACCGGCTTTGTAACCAAGCTTTAGAGGAGTTCTGCTG 656
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
Db 657 CCGACTAAGAGTACGCTGAGTATCCTCTCGAGGAGTTGATATATGAGGGCCATC 716
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240

```

Db      717  AACGTATGTCGACACATCCAGACATCTCAAGGGGACCTGTCCAGACGGCAGCGA 776
Qy      241  G1YGI5eranglYarSerTYrLYsArTYrPheserLupProglYasPhisProglY 260
Db      777  GGGGGGGCCCAAGCGGCGCAACCAAGAGACCTTCCCGACGAGAGATCAACCTGGG 836
Qy      261  MetLeuThsSerGlyLysArgSerHisLeuGlusSerSerArpProHis 277
Db      837  GTGTGGCCACTGGCAAGCGCTCACCTGAGATCTAGACAGACACCCAT 887

RESULT 10
MUSURKI      1810 bp  mRNA  linear  ROD 27-MAR-1997
LOCUS        Mus musculus uridine kinase mRNA, partial cds.
ACCESSION    U11783
VERSION       U11783.1 GI:471980
KEYWORDS     Mus musculus (house mouse)
SOURCE        Mus musculus
ORGANISM      Mus musculus (house mouse)
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        The functions and consensus motifs of nine types of peptide
              segments that form different types of nucleotide-binding sites
JOURNAL      Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE      94259063
PUBMED       8200357
REFERENCE    2 (bases 1 to 1810)
AUTHORS      Ropp, P.A. and Traut, T.W.
TITLE        Cloning and expression of a cDNA encoding uridine kinase from mouse
              brain
JOURNAL      Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE      97108719
PUBMED       8951040
FEATURES
    source     location/Qualifiers
              1..1810
              /organism="Mus musculus"
              /mol_type="mRNA"
              /db_xref="taxon:10090"
              /tissue_type="brain"
              <1..783
              /EC_number="2.7.1.48"
              /standard_name="uridine-cytidine kinase"
              /standard_start=1
              /product="uridine kinase"
              /protein_id="AAB50568.1"
              /db_xref="GI:471981"
              /translation="RROPRLGVSGGTASGKSTVCEKIMELGONEVDRQRKLYI
              LSGDCEYKVLTAQKAKLKGQYNFDPDAFNDLMKTKLNLVEGKTEVPTDFVT
              HSRLEPTTVVYPADVVLFGILVFTQELIDMHLFLFVTDSDVLSRVLADVORG
              RLEQILITQYTAIVKPAFEFECPTKKYADVILPRGDMVAINLIVQHIQDILNGDL
              CKHRRGPMGRNKRTPERBGDHPGVLATATKRSHLSSSRPH"
              join(37..63,412..426,451..465)
              /bound_moiety="ATP at catalytic site"
              misc_binding
              /bound_moiety="ATP at catalytic site"
              polyA_signal
              BASE COUNT 411 a 482 c 512 g 405 t
              ORIGIN
Alignment Scores:
Pred. No.:      3 97e-111      Length:      1810
Score:          1266.00      Matches:      241
Percent Similarity: 96.92%      Conservative: 11
Best Local Similarity: 92.69%      Mismatches: 8
Query Match:    87.31%      Indels:      0
DB:             10      Gaps:      0
US-09-896-522-2 (1-277) x MUSURKI (1-1810)
Qy      18  ArgProHisG1LysArgProPhenLeuIleG1ValSerG1YGIYThraLaseG1YLySer 37
Db      1  CGTCCAGCGCGCGCGGCGTTCCTCATCGCGGTAGCGCGGCAACCGCTAGTGCAGACTCA 60

```

```

Qy      38  ThrValCYeGluLysL1LemG1LysLeuLeuG1YGIYThraLaseG1YLySer 57
Db      61  ACAGTGTGAGAGATCAAGAGCTGTGGGACAAAGAGTGAACCGCGGAGCGG 120
Qy      58  LysValVal1LLeuSerG1nAspArgPheTYrLYsValLeuThraLag1nLysAla 77
Db      121  AAGTTGTATCTGTAGCCAGAGCTCTCTCAAGAGTTCTAGCGGCCAGAGAGGCC 180
Qy      78  LysAlaLeuLysG1YnTYrAnPheAspHisProAspAlaPheAspAnAspLeuMet 97
Db      181  AAGGCTTGAAGGAGACATCAATTTTGACCAACCCAGATGCTTTGATATGATCGATG 240
Qy      98  HisArgThreLysSerL1LeuValG1YGIYThraLaseG1YLySer 117
Db      241  CAGPAGACCTGAAACATGTGTGAAGCAAAAGTGAAGTCTCAAGATTTT 300
Qy      118  ValThHisSerArgLeuProG1nThraVal1LysProAlaAspVal1LLeuPhe 137
Db      301  GTGACCCACTCAAGTTACAGAGACCACTGTGTCTACCCAGCTGATGGTTCTGTT 360
Qy      138  GluG1YLeuValPheTYrSerG1nL1LysArgAspMetPheHisLeuArgLeuPhe 157
Db      361  GAGGCGATCTGGTATTCTACACCCAGAGATCCGGAGATGTTCCACCTGCGCTTT 420
Qy      158  ValAspThraSPSerAspValArgLeuSerArgArgValLeuArgAspValArgG1Y 177
Db      421  GTGGACACAGACTGTATGTAGCTGTCTCGAAGAGTTCTCGGAGATGTCAACAGAGA 480
Qy      178  ArgAspLeuG1nL1LeuThraG1nTYrThraPheValLysProAlaPheG1nL1 197
Db      481  AGGAGACTGAGAGATCTGACTCAGTACACCGCTTTGTGAACACAGCTTGAAGAG 540
Qy      198  PheCYeLeuProThraLysLYsTYrAlaAspVal1LLeuProArgG1YValAspAnMet 217
Db      541  TTCTGCTGCCCACTAAGAGTAAAGTGAAGCTGACGTATCATCTCGAGAGATTGATATG 600
Qy      218  ValAla1LLeuLysLeuValG1nHisL1LeuAsp1LLeuAnG1YLySer 237
Db      601  GTGGCCATCAACCTGATGTGCAACATCCAGAGATCTCAACGGGACCTGTCAAG 660
Qy      238  TrpHisArgG1YGIYSeranglYarSerTYrLYsArTYrPheserLupProglYasP 257
Db      661  CGGACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAT 720
Qy      258  HisProG1YMetLeuThsSerG1YLysArgSerHisLeuGlusSerSerArpProHis 277
Db      721  CACCGTGGGATGTGGCCACTGCGCAAGCGCTCACCTGAGTCTAGACAGACACCCAT 780

RESULT 11
BC015547      2072 bp  mRNA  linear  PRI 04-OCT-2001
LOCUS        Homo sapiens, Similar to uridine-cytidine kinase 1, clone MGC:9668
DEFINITION   IMAGE:3845821, mRNA, complete cds.
ACCESSION    BC015547
VERSION       BC015547.1 GI:15930229
KEYWORDS     MGC.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2072)
Strausberg,R.
Direct Substition
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
REMARK
COMMENT

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 20 Row: 1 Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES

SOURCE

1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 87..692
 /codon_start=1
 /product="similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="gi:15930230"
 /translation="MASAGGEDCESPAEADPHQPFLLIGVSGTASGKSYCEKIM
 ELIQNEVORQRKRVILSDRFKVLTAEBKAKALQGVNFDHDPADFNDLMRTIK
 NIVEGKTEVPTDYFTHSRILPETVVYPADVAFEGILFVYSQEIRDMFHLRLFVDT
 DSDVRLSRDKVRCRCHPRTSGOYGGHQPRAHPHSEM"

CDS

BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.78e-108 Length: 2072
 Score: 1239.00 Matches: 244
 Percent Similarity: 87.77% Conservatave: 0
 Best Local Similarity: 87.77% Mismatches: 1
 Query Match: 85.45% Indels: 33
 DB: 9 Gaps: 1

US-09-896-522-2 (1-277) x BC015547 (1-2072)

QY 1 Metlaseerlaaglygluaspdyguserproalaaprogiluiaaspargprohis 20
 Db 87 ATGGCTTCGGGGAGGAGGAGCTGCGAGGCCCGCGGAGCGGACCTGCCAC 146
 QY 21 Glnatgppheleuilegilyvaliserclgltthralasergilysserthvalcys 40
 Db 147 CAGCGGCCCTTCCTGATRGSGGTAGCGCGCGCACTGCCACCGGAAAGTCCAGCTGTGT 206
 QY 41 Glulysilemetguleuileugilyglnangluvalgulinatrginarglyvala 60
 Db 207 GAGAAATCATGAGATTGCTGGAAGAAAGAGGTGGAACGCGGACGCGAAGTGTGTC 266
 QY 61 Ileleuserglnasppargphetyrlyvalleuthralagluinlysalalysalaleu 80
 Db 267 ATCTCGGCCGAGCAGAGTTCTTAAGGTCTGACGCGAGGAGGAGGCGCAAGCCCTTG 326
 QY 81 Lysgilygintyrasppheasphileproasphaleasphaspaaspleamethisargthr 100
 Db 327 AAAAGACAGTACATTTTGAACATCCAGATCCCTTTGATATGATTTTGAATGACAGACT 386
 QY 101 Leuylasaniilevalgugilylserthvalglnuvalprothryraspphevalthrhis 120
 Db 387 CTGAAGAATCATCGTGAAGGCAAAAGGTGAGGTGCGACCTATGATTTTGTGACACAC 446
 QY 121 Sertrgleuprogiluthrthvalvaltyrproalaasppvalleulepneugilytyle 140
 Db 447 TCAAGGTTACAGAGACCAAGGTGTCTTCACTCGGACGTTGTTGAGGGGATC 506

QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 507 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTTTCGTGGACACC 566
 QY 161 AspSerAspValArgLeuSerArg-ArgValLeuArgAspValArgArgIleArgAspIle 180
 Db 567 GACTCCGACGTACAGCTGTCTCGAAGAG----- 594
 QY 180 uGluGlnIleLeuThrGlnTyrThrThrPheValYsProAlaPheGluGluPheCysIle 200
 Db 594 ----- 594
 QY 200 uProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAspMetValAlaI 220
 Db 595 ----ACAAAGATAGTACCGATGTGATGATCCACAGAGATGAGACAAATGTGTCAT 650
 QY 220 eAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisAr 240
 Db 651 CAACCTGATGTCAGACACATCCAGACATTTGAAATGTTGACATCTGCMAATGGCACCG 710
 QY 240 ggilygilyserannglyarserertyrlysarqthrpheserGluProGlyAspHisProG 260
 Db 711 AGAGGGGTCCAAAGGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCACTGG 770
 QY 260 yMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
 Db 771 GATGCTGACTCTGCGCAACGTCATATTGGAGTCCACAGACAGACCCAC 822

RESULT 12
 AC078885/c 254993 bp DNA linear HTG 15-MAY-2002
 LOCUS AC078885 chromosome 6 clone RP23-117123 strain C57BL/6J,
 DEFINITION WORKING DRAFT SEQUENCE, 44 unordered pieces.

ACCESSION AC078885 GI:18376842
 VERSION AC078885.8
 HTG; HTGS PHASE1; HTGS DRAFT.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.
 High Throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 254993)

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.
 Direct Submission
 Submitted (08-NOV-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Jan 26, 2002 this sequence version replaced gi:14488282.

COMMENT
 -----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://www.hpcg.org/sequence/mouse.html>
 Contact: hpcg@mednet.mgh.harvard.edu
 -----Summary Statistics

Center project name: AAP
 Sequencing vector: pUC18, L08752
 Chemistry: Dye-terminator Big Dye, 100%
 Consensus quality: 247096 at least Q20
 Consensus quality: 242518 at least Q30
 Consensus quality: 235589 at least Q40
 Estimated insert size: agarose-FP - N/A
 Estimated insert size: 254133 - sum-of-contigs
 Quality coverage: agarose-FP - N/A
 Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

 * NOTE: This is a 'working draft' sequence. It currently

* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 27415: contig of 27415 bp in length
* 27416 27435: gap of unknown length
* 27436 57817: contig of 30382 bp in length
* 57818 57837: gap of unknown length
* 57838 81637: contig of 23800 bp in length
* 81638 81657: gap of unknown length
* 81658 102155: contig of 20498 bp in length
* 102156 102176: gap of unknown length
* 102176 121200: contig of 19025 bp in length
* 121201 121220: gap of unknown length
* 121221 135073: contig of 13853 bp in length
* 135074 135093: gap of unknown length
* 135094 156689: contig of 21596 bp in length
* 156690 156710: gap of unknown length
* 156710 166185: contig of 9476 bp in length
* 166186 166205: gap of unknown length
* 166206 178798: contig of 12593 bp in length
* 178799 178818: gap of unknown length
* 178819 187247: contig of 8429 bp in length
* 187248 187267: gap of unknown length
* 187268 192492: contig of 5225 bp in length
* 192493 192512: gap of unknown length
* 192513 200084: contig of 7572 bp in length
* 200085 200104: gap of unknown length
* 200105 207054: contig of 6950 bp in length
* 207055 207074: gap of unknown length
* 207075 212944: contig of 5870 bp in length
* 212945 212964: gap of unknown length
* 212965 219468: contig of 6504 bp in length
* 219469 219488: gap of unknown length
* 219489 220369: contig of 881 bp in length
* 220370 220389: gap of unknown length
* 220390 221164: contig of 775 bp in length
* 221165 221184: gap of unknown length
* 221185 222796: contig of 1612 bp in length
* 222797 222816: gap of unknown length
* 222817 224762: contig of 1946 bp in length
* 224763 224782: gap of unknown length
* 224783 225999: contig of 1217 bp in length
* 226000 226019: gap of unknown length
* 226020 227409: contig of 1390 bp in length
* 227410 227429: gap of unknown length
* 227430 230145: contig of 2716 bp in length
* 230146 230165: gap of unknown length
* 230166 231842: contig of 1677 bp in length
* 231843 231862: gap of unknown length
* 231863 232786: contig of 924 bp in length
* 232787 232806: gap of unknown length
* 232807 233137: contig of 331 bp in length
* 233138 233157: gap of unknown length
* 233158 234519: contig of 1362 bp in length
* 234520 234539: gap of unknown length
* 234540 236416: contig of 1877 bp in length
* 236417 236436: gap of unknown length
* 236437 237967: contig of 1531 bp in length
* 237968 237987: gap of unknown length
* 237988 239220: contig of 1233 bp in length
* 239221 239240: gap of unknown length
* 239241 239367: contig of 127 bp in length
* 239368 239387: gap of unknown length
* 239388 239959: contig of 572 bp in length
* 239960 239979: gap of unknown length
* 239980 241016: contig of 1037 bp in length
* 241017 241036: gap of unknown length
* 241037 242391: contig of 1355 bp in length
* 242392 242411: gap of unknown length

242412 243823: contig of 1412 bp in length
* 243824 243843: gap of unknown length
* 243844 245227: contig of 1384 bp in length
* 245228 245247: gap of unknown length
* 245248 246354: contig of 1087 bp in length
* 246355 246374: gap of unknown length
* 246375 247451: contig of 1097 bp in length
* 247452 247471: gap of unknown length
* 247472 248780: contig of 1309 bp in length
* 248781 248800: gap of unknown length
* 248801 250041: contig of 1241 bp in length
* 250042 250061: gap of unknown length
* 250062 251173: contig of 1112 bp in length
* 251174 251193: gap of unknown length
* 251194 252713: contig of 1520 bp in length
* 252714 252733: gap of unknown length
* 252734 254014: contig of 1281 bp in length
* 254015 254034: gap of unknown length
* 254035 254176: contig of 142 bp in length
* 254177 254196: gap of unknown length
* 254197 254993: contig of 797 bp in length.

FEATURES
source
1. 254993
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-117123"
/sex="male"
1. 27415
/note="assembly_name:Contig138"
27436..57817
/note="assembly_name:Contig137"
clone_end:5p6
vector_side:left"
57838..81637
/note="assembly_name:Contig136"
81658..102155
/note="assembly_name:Contig135"
102176..121200
/note="assembly_name:Contig134"
121221..135073
/note="assembly_name:Contig133"
135094..156689
/note="assembly_name:Contig132"
156710..166185
/note="assembly_name:Contig131"
clone_end:T7
vector_side:left"
166206..178798
/note="assembly_name:Contig130"
178819..187247
/note="assembly_name:Contig129"
187268..192492
/note="assembly_name:Contig128"
192513..200084
/note="assembly_name:Contig127"
200105..207054
/note="assembly_name:Contig126"
207075..212944
/note="assembly_name:Contig125"
212965..219468
/note="assembly_name:Contig124"
219489..220369
/note="assembly_name:Contig123"
220390..221164
/note="assembly_name:Contig122"
221185..222796
/note="assembly_name:Contig121"
222817..224762
/note="assembly_name:Contig120"
224783..225999


```

/misc_feature      /note="assembly_name:Contig119"
                   226020..227409
/misc_feature      /note="assembly_name:Contig118"
                   227430..230145
/misc_feature      /note="assembly_name:Contig117"
                   230166..231842
/misc_feature      /note="assembly_name:Contig116"
                   231863..232786
/misc_feature      /note="assembly_name:Contig115"
                   232807..233137
/misc_feature      /note="assembly_name:Contig114"

```

Alignment Scores:

```

Pred. No.:      3,2e-103      Length:      254993
Score:          1212.00      Matches:      230
Percent Similarity: 97.58%      Conservative: 12
Best Local Similarity: 92.74%      Mismatches: 6
Query Match:      83.59%      Indels: 0
DB:               2          Gaps: 0

```

US-09-896-522-2 (1-277) x AC078885 (1-254993)

```

QY      30 GlycylThrAlaSerGlyLysSerThrValCysGluLysIleMetGluLeuGlyGln 49
DB      156682 GGCAGGACCCGCTAGTGGCAAGTCAACAGTGTGAGAAGATCATGAGAGCTGGACAG 156623

QY      50 AsnGluValGluGlnArgGlnArgLysValIleLeuSerGlnAspArgPheTyrLys 69
DB      156622 AACCAAGTGAAGCCCGGACGCGACGAGTGTGATCCCGACCCAGAGTGGCTTCTACAG 156563

QY      70 ValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisPro 89
DB      156562 GTTCTGACGGCCGACGAGAACCCAGCTTTGAAGGACAGTACATTTTGACACCA 156503

QY      90 AspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGluGlyLysThr 109
DB      156502 GATGCTTTTGTATATGATCTGATCACAAGACCCCTGAAAAACATTTGTGAAGGCACAAACT 156443

QY      110 ValGluValProThrTyrAspPheValThrHisSerArgLeuProGluThrThrValVal 129
DB      156442 GTGAGGTCCTTACTTATGATTTGTGACCACTCAAGATTACCAAGAGACATGTGTC 156383

QY      130 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGlnIleArg 149
DB      156382 TACCACCTGATGTGTTCTGTTGAGGAGCATCTTGGATTCTACACCCGGAGATCCGG 156323

QY      150 AspMetPheHisLeuAlaGluPheValAspThrAspSerAspValAlaGluSerArgArg 169
DB      156322 GACATGTTCACCTGCGCCCTCTTGTGTGACACAGACTCTGATTTAGGCTGCTCTCGAAGA 156263

QY      170 ValLeuArgAspValAlaGArgGlyAlaArgAspLeuGluGlnIleLeuThrGlnTyrThrThr 189
DB      156262 GTTCTCGGAGTGTGCAACGAGAGGAGCCTGGAACAGATCTGACTCAGTATACCGCC 156203

QY      190 PheValLysProAlaPheGluGlnPheCysLeuProThrLysLysTyrAlaAspValIle 209
DB      156202 TTTTGAACAACAGCCTTTGAGGAGTTCGCTGCGCATGAAGAGTACOGCTGAGTATC 156143

QY      210 IleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGlnHisIleGlnAsp 229
DB      156142 ATCCCTCGAGAGTGTGATATATGATGCGCATCACTGATCCGCAACATCCAGAGAC 156083

QY      230 IleLeuAsnGlyAspIleCysLysValPheHisArgGlyGlySerAsnGlyArgSerTyrLys 249
DB      156082 ATCTCTCAACGGGAGCTGTGTGAACAGACCGAGCGCGGCCCAACGGCGCAACAG 156023

QY      250 ArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerGlyLysArgSerHis 269
DB      156022 AGGACCTTCCCGGACGACAGAGATCACTCGGGGTGTGGCCATGCGCAACCGCTCACAC 155963

QY      270 LeuGluSerSerArgProHis 277
DB      155962 CTGGAGTCTAGCAGACAGACCCCAT 155939

```

```

RESULT 13
MUSURKIA          952 bp      DNA      linear      ROD 05-MAR-2001
LOCUS             Mus musculus uridine kinase pseudogene.
DEFINITION        L31784
ACCESSION         L31784.1 GI:471982
VERSION           pseudogene: uridine kinase.
KEYWORDS          Mus musculus (house mouse).
SOURCE            Mus musculus
ORGANISM          Mus musculus
REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE             1 (bases 1 to 952)
AUTHORS           Traut,T.W.
TITLE             The functions and consensus motifs of nine types of peptide
JOURNAL            segments that form different types of nucleotide-binding sites
MEDLINE           Eur. J. Biochem. 222 (1), 9-19 (1994)
PUBMED            94259063
PUBMED            8200357
REFERENCE          2 (bases 1 to 952)
AUTHORS           Ropp,P.A. and Traut,T.W.
TITLE             Cloning and expression of a cDNA encoding uridine kinase from mouse
JOURNAL            brain
MEDLINE           Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
PUBMED            97108719
PUBMED            8951040
FEATURES           source
                   1..952      Location/Qualifiers
                   /organism="Mus musculus"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:10090"
                   /tissue_type="kidney"
                   76..913      /standard_name="uridine-cytidine kinase"
                   /pseudo
                   /codon_start=1
                   /product="uridine kinase"
                   misc_difference 130
BASE COUNT        230 a      231 c      277 g      214 t
ORIGIN
1 MetAlaSerAlaGlyGlyLeuAspCysGluSerProAla----ProGluAlaAspArgPr 19
DB      76 ATGGCTTCGGCGGAGGCGGTGCTCAGAGAGTACGCGCCGCGAGCGCATCTCC 135
QY      19 ohisGlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrVal 39
DB      136 TCAGCTTGCGCCCTTCTCATCGGCGGTGAGTGGCGGACCGCTAGTGGCAAGTCAACAGT 195
QY      39 lCysGluLysIleMetGluLeuLeuGlyGlnAsnGlyValGluGlnArgGlnArgLysVal 59
DB      196 GTGTGAGGAATCATGAGCTGCTGGGACAGAAATGAATGAGACCTCCGTTAGCGCAAGTT 255
QY      59 lValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAl 79
DB      256 GGTGATCTCGAGCGACAGACTGCTTACAAAGTTCTGACGGCTGAGACAGAAAGCCAAAGC 315
QY      79 AleuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99
DB      316 TTGGAAGGACAGTACAAATTTTGACCATGACAGATGCTTTTGTATTAATGATCTGTGACAA 375
QY      99 gThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValTh 119

```


US-09-896-522-2 (1-277) x AC118474 (1-192336)

```

Qy      1 MetAlaserAlaglygylgluaspCygluSerProAla-----ProGluAlaAspArgPr 19
Db      77706 ATGGCTTGGCGGAGGCGGTGGCTCAAGAGTACTGCGCTGCCGAGGCGCATCTTCC 77765
Qy      19 ohIsgInArPrPheLeuIleGIyValSerGIyGlyThrAlaSerGIyLysSerThrVa 39
Db      77766 TCAGCCTTGGCGGCTTCTCATCGGCTGAGTGGCGGCAACCTTGTGCAAGTCAACAGT 77825
Qy      39 lCyGluLysIleMeCgluLeuLeuGIyGlnAsnGluValGluGlnArgGlnArgLysVa 59
Db      77826 GTGCGAGAAGATCATGAGCTGCTGGGACAAGATGATGACCTCCGCTGAGCCCAAGTT 77885
Qy      59 lValIleLeuSerGIlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAl 79
Db      77886 GGTCATCTTGAGCCAGGAGCTGCTTCTACAGGTTCTGAGGCTGAGCAAGAAAGCCAAAGGC 77945
Qy      79 aLeuLysGIyGlnTyraSnPheAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99
Db      77946 TTGAGAGGACAGATACATTTTGACCATGACAGATGCTTTGATATGATCTGATGACAA 78005
Qy      99 gThLeuLysAsnIleValGluGIyLysThrValGluValProThrTyraSpPheValTh 119
Db      78006 GACCTTGAAGAAACATTATGAAAGCAAGTTGTTAGGTTCTTACCTATGATTTTGTGAC 78065
Qy      119 rHisSerArgLeuProGluThrThrValValTyraProAlaAspValValLeuPheGlu 139
Db      78066 CCACCTACAGTTTACACAGTACCACTGTGGTCTACCAATTGAGCGGTTCTGTTCGAGGG 78125
Qy      139 yIleLeuValPheTySerGIlnGluIleArgAspMetPheHisLeuArgLeuPheValAs 159
Db      78126 CATCTTGATATTCTACACTCAAGAGATCCGGGACATGTTTCACTCGCGCTTTGTGGA 78185
Qy      159 pThAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGIyLysArgAs 179
Db      78186 CACAGACCCTGATGTAGGCTGTCTTGAAAGATTCTCGGATGTGCATAGAGAAAGGGA 78245
Qy      179 pLeuGIlnGlnIleLeuThrGIlnTyraThrPheValLysProAlaPheGIlnGluPheCy 199
Db      78246 CTTGAGAGCAATCTGACTGCTGACGCCACCACTTGTGAATACGCGCTTCGAAAGATTCTG 78305
Qy      199 sLeuProThrLysTyraLAspValIleIleProArgGIyValAspAsnMetValAl 219
Db      78306 CCTGCCGACTTA-----GAAGTGATCATCCCTCGAGGCTTGACATATGCTGGC 78355
Qy      219 aIleAsnLeuIleValGlnHisIleGIlnAspIleLeuAsnGIyAspIleCysLysThrPh 239
Db      78356 CATCAACCTGATCGTATTAACAATCCAGACATCTCTCAACCGGGAACCTGTGCAAGCGCA 78415
Qy      239 sArgGIyLysSerArgGIyArgSerTyraArgThrPheSerGIlnProGIyAspHisAr 259
Db      78416 CCGAGAGAGGCCCAATGGGACAGCCACACAGAGATCTTACCAGATGACAGAGATCAACC 78475
Qy      259 oGIyMeIeuThrSerGIyLysArgSerHisLeuGlnSerSerSerArgProHis 277
Db      78476 TGGGGTGTGGCCACTGTTAAACCTGACACTGGAAGTCTAGCAGCAGATCCCAT 78530

```

RESULT 15

AC076974 211075 bp DNA linear ROD 16-NOV-2002
 LOCUS Mus musculus chromosome 8 clone rp23-267j18 strain C57BL/6J,
 DEFINITION complete sequence.
 AC076974 AC076974 GI:25046382
 VERSION AC076974.33
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 211075)
 AUTHORS Deschamps,S., Gu,W. and Roe,B.A.

```

TITLE      Mus musculus Chromosome 1 BAC Clone rp23-267j18
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 211075)
AUTHORS    Deschamps,S., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL    Submitted (01-AUG-2000) Department Of Chemistry And Biochemistry,
           The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
           OK 73019, USA
REFERENCE  3 (bases 1 to 211075)
AUTHORS    Deschamps,S., Be, Zemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL    Submitted (12-JUL-2002) Department Of Chemistry And Biochemistry,
           The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
           OK 73019, USA
REFERENCE  4 (bases 1 to 211075)
AUTHORS    Deschamps,S., Be, Zemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL    Submitted (18-JUL-2002) Department Of Chemistry And Biochemistry,
           The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
           OK 73019, USA
REFERENCE  5 (bases 1 to 211075)
AUTHORS    Deschamps,S., Be, Zemmer,B., Gu,W. and Roe,B.A.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2002) Department Of Chemistry And Biochemistry,
           The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
           OK 73019, USA
COMMENT    On Nov 16, 2002 this sequence version replaced gi:21734996.
           ----- Genome Center
           Center: Department Of Chemistry And Biochemistry
           The University Of Oklahoma
           Center code:UOKNOR
           -----
FEATURES   source
           1..211075
              location/Qualifiers
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /strain="C57BL/6J"
              /db_xref="taxon:10090"
              /chromosome="8"
              /clone="rp23-267j18"
              /clone_lib="RPCT - 23 Female (C57BL/6J) Mouse BAC Library"
              /note="This is one of two clones in the same well from
                    rp23-267j18"
BASE COUNT 59757 a 44080 c 44532 g 62706 t
ORIGIN
Alignment Scores:
Pred. No.: 2,376-90 Length: 211075
Score: 1076.00 Matches: 225
Percent Similarity: 85.30% Conservative: 13
Best Local Similarity: 80.65% Mismatches: 36
Query Match: 74,21% Indels: 6
DB: 10 Gaps: 2

```

US-09-896-522-2 (1-277) x AC076974 (1-211075)

```

Qy      1 MetAlaserAlaglygylgluaspCygluSerProAla-----ProGluAlaAspArgPr 19
Db      195148 ATGGCTTGGCGGAGGCGGTGGCTCAAGAGTACTGCGCTGCCGAGGCGCATCTTCC 195207
Qy      19 ohIsgInArPrPheLeuIleGIyValSerGIyGlyThrAlaSerGIyLysSerThrVa 39
Db      195208 TCAGCCTTGGCGGCTTCTCATCGGCTGAGTGGCGGCAACCTTGTGCAAGTCAACAGT 195267
Qy      39 lCyGluLysIleMeCgluLeuLeuGIyGlnAsnGluValGluGlnArgGlnArgLysVa 59
Db      195268 GTGCGAGAAGATCATGAGCTGCTGGGACAAGATGATGACCTCCGCTGAGCCCAAGTT 195327
Qy      59 lValIleLeuSerGIlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAl 79
Db      195328 GGTCATCTTGAGCCAGGAGCTGCTTCTACAGGTTCTGAGGCTGAGCAAGAAAGCCAAAGC 195387
Qy      79 aLeuLysGIyGlnTyraSnPheAspHisProAspAlaPheAspAsnAspLeuMetHisAr 99

```

```

Db      195388 TTGAAAGGACAGATCAATTTTGACCATGCAGATGCTTTGATATGATCGATGACAA 195447
QY      99 gThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValTh 119
Db      195448 GACCTTGAAAACATTTATGTGAAGCAAGGTAAGTTTCTTAACTTATGATTTTGTGAC 195507
QY      119 rHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlu 139
Db      195508 CCACTCAAGGTTACCACTGACCTGCTGCTTACACAATTGATCGGTTCTGTTGAGGG 195567
QY      139 YIleLeuValPheTyrSerGluGluIleArgAspMetPheHisLeuArgLeuPheValAs 159
Db      195568 CATCTTGATTTCTACACTCAAGAGATCCGGACATGTTTCACTGCGCTTGTGGA 195627
QY      159 pThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAs 179
Db      195628 CACAGACCCTGATGTGTAGGCTGTCTTGAAAGAGTTCTTGAGATGTGCAATGAGGAAAGGA 195687
QY      179 pLeuGluGluIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCy 199
Db      195688 CTTGAGCAATCTCTGACTGACGACCACTGATTTGAAATCGGCTTCAAGAGTTCTG 195747
QY      199 sLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAl 219
Db      195748 CTGCGCACTTA-----GAAGTGCATCATCCCTCGAGGGCTTGACAAATATGTGCG 195797
QY      219 aIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHis 239
Db      195798 CATCAACCTGATGTATTAACACATCCAGACATCTCAACGGGACCTGTGCAAGCGGCA 195857
QY      239 sArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisPr 259
Db      195858 CCGAGAGGGGCCCAATGGGCAACGCCACAAAGAGACCTTACCCGAGTCAGAGATCACCC 195917
QY      259 OGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
Db      195918 TGGGGTGTGTGGCCACTGTAAACGCTGACACCTGAGTCTAGCAGCAGATCCCAT 195972

```

Search completed: November 25, 2003, 09:29:17
 Job time : 4279 secs

This Page Blank (uspto)


```

QY 89 CCGAGATGCTTGGCGGAGGCGAGAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 148
D 90 CCGAGATGCTTGGCGGAGGCGAGAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 149
QY 149 CGCAGCAGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGCGCATGCGAGCGGAGGTCGACCG 208
D 150 CGCAGCAGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGCGCATGCGAGCGGAGGTCGACCG 209
QY 209 TGTGTGAGGAAGATCATGAGAGTTGCTGAGAGCAAGAGGTTGAGAGGCGGAGG 268
D 210 TGTGTGAGGAAGATCATGAGAGTTGCTGAGAGCAAGAGGTTGAGAGGCGGAGG 269
QY 269 TGTGTGAGGAAGATCATGAGAGTTGCTGAGAGCAAGAGGTTGAGAGGCGGAGG 328
D 270 TGTGTGAGGAAGATCATGAGAGTTGCTGAGAGCAAGAGGTTGAGAGGCGGAGG 329
QY 329 CTTGGAAGAGCAATGATTTTGAATCCATCCAGATGCTTGTGATGATTTGATGACA 388
D 330 CTTGGAAGAGCAATGATTTTGAATCCATCCAGATGCTTGTGATGATTTGATGACA 389
QY 389 GGAATCTGGAAGATCATGAGAGGCGGAGAGGCGGAGGCGGAGGCGGAGG 448
D 390 GGAATCTGGAAGATCATGAGAGGCGGAGAGGCGGAGGCGGAGGCGGAGG 449
QY 449 CACATCAAGGTTTCAAGAGCAAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 508
D 450 CACATCAAGGTTTCAAGAGCAAGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 509
QY 509 GCAATCTGATGTTTCTAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 568
D 510 GCAATCTGATGTTTCTAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 569
QY 569 ACACCGAGCTCCGAGCTGAGGCTGCTGAGAGAGTTCTCCGAGAGGAGGAGG 628
D 570 ACACCGAGCTCCGAGCTGAGGCTGCTGAGAGAGTTCTCCGAGAGGAGGAGG 629
QY 629 ACCTGAGGAGATCTGAGCGGAGTACCACTTCTGAGAGGCGGAGGAGGAGG 688
D 630 ACCTGAGGAGATCTGAGCGGAGTACCACTTCTGAGAGGCGGAGGAGGAGG 689
QY 689 GCCTGCGGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 748
D 690 GCCTGCGGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 749
QY 749 CCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 808
D 750 CCATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
QY 809 ACCGAGAGGAGTCCATGAGGAGGAGTACCAAGCGGAGCTTCTGAGAGGAGG 868
D 810 ACCGAGAGGAGTCCATGAGGAGGAGTACCAAGCGGAGCTTCTGAGAGGAGG 869
QY 869 CTGGAGATGCTGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 928
D 870 CTGGAGATGCTGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 929
QY 929 GGGGCTGCGGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 988
D 930 GGGGCTGCGGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 989
QY 989 GGGGCGGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 1013
D 990 GGGGCGGAGCTTCTGAGAGGAGTACCAAGCGGAGCTTCTGAGAGGAG 1014

```

RESULT 2

```

US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase

```

```

; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1

Query Match      51.3%; Score 832.4; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 1.8e-212;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGCTTCCGCGGAGGAGGCGAGAGCTGCGAGAGCCCGCGCGGAGGCGGAGCGTCCGAC 153
D 1 ATGGCTTCCGCGGAGGAGGCGAGAGCTGCGAGAGCCCGCGCGGAGGCGGAGCGTCCGAC 60
QY 154 CAGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGAGCTGCGAGCGGAGGTCGACCGTGT 213
D 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGGCGGAGCTGCGAGCGGAGGTCGACCGTGT 120
QY 214 GAGAAATCATGAGAGTTGCTGAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 273
D 121 GAGAAATCATGAGAGTTGCTGAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 274 ATCTGAGCGAGGAGGAGGTTCTCAAGGCTCTGAGCGGAGGAGGAGGAGGAGGAGG 333
D 181 ATCTGAGCGAGGAGGAGGTTCTCAAGGCTCTGAGCGGAGGAGGAGGAGGAGGAGG 240
QY 334 AAGAGCAGTACATTTTGAATCCATCCAGATGCTTGTGATGATTTGATGACAGAGCT 393
D 241 AAGAGCAGTACATTTTGAATCCATCCAGATGCTTGTGATGATTTGATGACAGAGCT 300
QY 394 CTGAAAGATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453
D 301 CTGAAAGATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 454 TCAAGGTTTCAAGAGCAAGGAGGAGTCTCAAGGCTCTGAGAGGAGGAGGAGGAGG 513
D 361 TCAAGGTTTCAAGAGCAAGGAGGAGTCTCAAGGCTCTGAGAGGAGGAGGAGGAGG 420
QY 514 TTGATGTTTCAAGCGGAGGAGATCCGAGCATGTTTCACTTGAGGCTTCTGAGACCC 573
D 421 TTGATGTTTCAAGCGGAGGAGATCCGAGCATGTTTCACTTGAGGCTTCTGAGACCC 480
QY 574 GATCCGAGCTGAGGCTGCTCTGAGAGAGTTTCTCCGAGAGCTGCGGAGGAGGAGG 633
D 481 GATCCGAGCTGAGGCTGCTCTGAGAGAGTTTCTCCGAGAGCTGCGGAGGAGGAGG 540
QY 634 GAGCAGATCTGAGCGGAGTACCACTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 693
D 541 GAGCAGATCTGAGCGGAGTACCACTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 694 CCGAGCAAGAGATGATGAGTATGATCCAGAGAGTGTGACAAATATGTTGTCATC 753
D 601 CCGAGCAAGAGATGATGAGTATGATCCAGAGAGTGTGACAAATATGTTGTCATC 660
QY 754 AACCTGATGTTGAGAGCAATCCAGAGCATTTGAAATGTTGATCATCTGCAAAATG 813
D 661 AACCTGATGTTGAGAGCAATCCAGAGCATTTGAAATGTTGATCATCTGCAAAATG 720
QY 814 GAGAGGTTCAATGAGGAGGAGTCAAGCGGAGCTTCTTCTGAGGAGGAGGAGGAGG 873
D 721 GAGAGGTTCAATGAGGAGGAGTCAAGCGGAGCTTCTTCTGAGGAGGAGGAGGAGG 780
QY 874 ATGCTGAGCTTCTGAGGAGGAGTCAATTTGAGTGTGAGTGTGAGGAGGAGGAGGAGG 927
D 781 ATGCTGAGCTTCTGAGGAGGAGTCAATTTGAGTGTGAGTGTGAGGAGGAGGAGGAGG 834

```

RESULT 3

US-09-221-017B-368
Sequence 368, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: ROSS, BRUCE C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA: P22911
FILING DATE: 09-APR-1998
APPLICATION NUMBER: PCT/AU98/01023
APPLICATION DATA:
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, GLADYS H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 5687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...5687
US-09-221-017B-368

Query Match 6.6%; Score 106.8; DB 4; Length 5687;
Best Local Similarity 54.3%; Pred. No. 1.6e-18;
Matches 238; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 341 AGTACATTTTGGACATCCAGATGCTTGTATGATTTGATGACAGGACTCTGAAGA 400
DB 4854 AGCTGAACACTGACACATCCATGCTTTGACACAGGATATGTCGTCAGGATATTTCTC 4913
QY 401 ACATCGTGAGGGCAAAACGGTGAGGTGCCGACTATGATTTTGTGACACACTCAAGT 460
DB 4914 CTCTGAAGGAGGCAAAACGATAGAGGTCGGGTCTATTCTTTCGTAGACCAATCGTT 4973

QY 461 TACGAGACACACGNGTCTACCCCTGCGAGCTGTTCTGTTGAGGCATCTTGCTG 520
DB 4974 TCGAAGAAAAAGTACACCGTTTCCTGCGCAAAAGTATGCTAGTGAATCTGATAT 5033
QY 521 TCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACCGACTCG 580
DB 5034 TCGAAGAACAAAGCTGCGGGATCTGATGATGTGAATATTTCTGTCGATCCGATCGCG 5093
QY 581 ACGTGAGCTGTCTGAGAGTTCTCGGAGGT--GCGCGAGGAGGACCTTGAGC 637
DB 5094 ATATTCGATTGGCGCGCCCTTGTCGCGATGTCAGGAGCGGGACCAATATGATAT 5153
QY 638 AGATTCTGACGAGTACACACCTTCGTGAGCGCGCTTCGAGGAGTCTGCTCCGA 697
DB 5154 CGGTATTGCAACAATCTTCAAGTCCGCTTCGATGACAGAGATTTTGTGGAACAT 5213
QY 698 CAAAGAGATGCGCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATCAAC 757
DB 5214 CCAAGCGTACCGCATCTGATCATTCGGAAGGTGATTCATTCGTTGCGCTCTCAC 5273
QY 758 TGATCGTGAAGACATCC 775
DB 5274 TCTTGTCCAAAAAATCC 5291

RESULT 4

US-08-961-527-76/c
Sequence 76, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
OPERATING SYSTEM: HP Vectra 486/33
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 1001 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-76

Query Match 6.3%; Score 101.8; DB 4; Length 1001;
Best Local Similarity 53.2%; Pred. No. 4.5e-17;
Matches 239; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

QY 345 CAATTTTGACATCCAGATGCTTGTATGATTTGATGACAGGACTCTGAAGAACAT 404

```
Db 10001 CAACCTATGACCATCTTTTGGCTTGTATGACAGCTTGATGATGACGACATTTAGCAATT 9942
Qy 405 CGTGGAGGGCAAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACACTCAAGTTACC 464
Db 9941 GTTGGACAGGGGCTCCGGTGGACATCCGACCTACGATATACAGAGATACAGAGATAG 9882
Qy 465 AGAGAACCAAGGTGTCTACCTCGGAGAGTGTTCTGTTTGAAGGCATCTTGGTGTCTA 524
Db 9881 CAAGACCTATGTCGAGGAACCTCAAGATGCTTTATGTTGAGGGTATTTTGGCTTGA 9822
Qy 525 CAGCAGAGATATCCGGGACATGTTCCACCTGCGCTCTGTCGAGACCCGACCTCGACGT 584
Db 9821 GGAACAAGCTGTGGCGCATTTGATGATATCAAGATTTTGTGATACGATGACGATGT 9762
Qy 585 CAGGCTGTCTGGAAGAGTCTCCGGGA---CGTCCGCGAGAGGAGGAGACCTGAGCAGAT 641
Db 9761 GCGCATATTCTGTGATCAAGCGTATGAGAGAGCGTGGCCGTAGCCTTGATAGCGT 9702
Qy 642 TCTGACCGCATACACACCTTCTGTGMAAGCCGCTTGAGAGAGTTTCTGCCCGACAAA 701
Db 9701 TATTAAACCACTAGTAAAGTGTGTCMAACCAATGATACCAAGTTTATGAGTCACTAA 9642
Qy 702 GAAGTATCCGATGATGATCATCCACGAGAGTGGAACAATATGTTGTCATCACTGAT 761
Db 9641 GCGTATGCTGATATCGTCAATCTCTGAAGGGTTAGCAATACGTTGCTATGACTGTT 9582
Qy 762 CGTGCAGCAATCCAGACATCTGATG 790
Db 9581 GACGACCAAGATGCAAAAGATTTGGAAG 9553
```

```
RESULT 5
US-09-134-001C-781
; Sequence 781, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 781
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-781
```

```
Query Match 5.3%; Score 86.2; DB 4; Length 651;
Best Local Similarity 51.5%; Pred. No. 18-13;
Matches 224; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

Qy 346 AATTGATGACATCCAGATGCTTTGATATGATTTGATGACAGACTCTGAGAGATC 405
Db 196 AATTATGACCAACCATTTGATTTGATATGATTTACTATTCAATTTAAAGATTTA 255
Qy 406 GTGAGGGCAAAAACGCTGAGAGTCCGACCTATGATTTTGTGAACACTCAAGTTACCA 465
Db 256 AGAATGAAAAACAGTAGAGTACTACGATATATTCTCAACATACACGATAGTAAA 315
Qy 466 GAGACCAAGGTGTCTACCTCGGAGAGTGTCTGTTGAGGGACCTGGGTGTCTAC 525
Db 316 GAAACAAATTGATTTGATTCAAAAGATGTATTTCGTAGAGGTATCTTTGCTTTGAA 375
Qy 526 AGCCAGAGATCCGAGCATGTTCCACTGCGCTCTTGTGAGACACCGACTCCGAGCTC 585
Db 376 AACAACAATTAAGACATGATGATGTGAAATTTATGTCGATCTGATGATGATTTA 435
```

```
Qy 586 AGCTGTCTGGAAGAGTTCTCCGGGA---CGTCCGCGAGGAGGAGGACCTGAGCAGATT 642
Db 436 CGAATTTTACGTAGGGCTTACAAAGAGCTGATGTAAGAGCGTGTGTAACAAATGATGTA 495
Qy 643 CTGACGAGATACACCACTTCTGTGAAGCCGCGCTTCGAGAGAGTTCTGCTGCGCAAAAG 702
Db 496 ATTATCAATATCTTTAATGATAGACCTTAGACATGACCAATTTATGAACTTACAAA 555
Qy 703 AAGTATCCGATGATATATCCACGAGAGTGGAACAATATGTTGCCATCAACTGATC 762
Db 556 AAGCATGTCGACATCAATATTTCTGAAGAGGAGCAGCAATTAAGTTCCATATATTATG 615
Qy 763 GTGACGACATCCAG 777
Db 616 ACTACTAAATCCAG 630
```

```
RESULT 6
US-09-107-532A-1370
; Sequence 1370, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denek
; REGISTRATION NUMBER: 40,489
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: circular
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1..633
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
US-09-107-532A-1370

Query Match 4.5%; Score 73.4; DB 4; Length 633;
Best Local Similarity 50.1%; Pred. No. 4,6e-10;
```



```

1 NAME/KEY: misc_feature
2 LOCATION: (120001)..(135000)
3 OTHER INFORMATION: n=a or c or g or t
4 NAME/KEY: misc_feature
5 LOCATION: (135001)..(150000)
6 OTHER INFORMATION: n=a or c or g or t
7 NAME/KEY: misc_feature
8 LOCATION: (150001)..(165000)
9 OTHER INFORMATION: n=a or c or g or t
10 NAME/KEY: misc_feature
11 LOCATION: (165001)..(180000)
12 OTHER INFORMATION: n=a or c or g or t
13 NAME/KEY: misc_feature
14 LOCATION: (180001)..(195000)
15 OTHER INFORMATION: n=a or c or g or t
16 NAME/KEY: misc_feature
17 LOCATION: (195001)..(210000)
18 OTHER INFORMATION: n=a or c or g or t
19 NAME/KEY: misc_feature
20 LOCATION: (210001)..(225000)
21 OTHER INFORMATION: n=a or c or g or t
22 NAME/KEY: misc_feature
23 LOCATION: (225001)..(240000)
24 OTHER INFORMATION: n=a or c or g or t
25 NAME/KEY: misc_feature
26 LOCATION: (240001)..(255000)
27 OTHER INFORMATION: n=a or c or g or t
28 NAME/KEY: misc_feature
29 LOCATION: (255001)..(270000)

```

LOCATION: (270001) .. (285000)

```

;
; OTHER INFORMATION: n=a or c
; NAME/KEY: misc_feature
;

```

; LOCATION: (285001) .. (300000)

OTHER INFORMATION: n=a or c

```
; NAME/KEY: misc_feature
```

LOCATION: (300001) .. (315000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

LOCATION: (315001) .. (330000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

LOCATION: (330001) - (345000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

INAME/ID1: IIIIBC_ICACUIC
LOCATION: (34EN01) /3600000

OTHER INFORMATION: 2-2 OF 3

OTHER INFORMATION: IED ON C
NAME/KEY: MISC FEATURES

```
NAME/KEY: IIIIBC_LeadCute
LOCATION: (3360001) (3375000)
```

LOCATION: (360001) .. (3/5000/

OTHER INFORMATION: H=4 OF C
 11/11/77 11:11 AM EST

```
; NAME/key: misc_feature
```

LOCATION: (375001) .. (390000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc_feature

LOCATION: (390001)..(405000)

OTHER INFORMATION: n=a or c

```
NAME/KEY: misc_feature
```

LOCATION: (405001) .. (420000) ;

OTHER INFORMATION: n=a or c

NAME/KEY: misc_feature;

LOCATION: (420001) .. (435000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

LOCATION: (435001) .. (450000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

LOCATION: (450001) - (465000)

OTHER INFORMATION: n=a or c

NAME/KEY: misc feature

LOCATION: (465001) (480000)

OTHER INFORMATION: n-3 or c

NAME/KEY, misc feature

NAME/KEY: LIBC_REALTIME;

```

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature

Query Match      3.4%; Score 55.8; DB 4; Length 1230025;
Best Local Similarity 49.9%; Pred. No. 0.00083;
Matches 197; Conservative 0; Mismatches 192; Indels 6; Gaps 2;

QY      341 AGTACAAATTTGACCATCCAGATGCCCTTGATTAATGATTTGATGCACAGGACTCTGAAGA 400
Db      835626 ATTTAATTTGGGATCATCCGAGCCGCTTGATTAATGACTTATTAATTCAGACTAABAAC 835567
QY      401 ACATCGTGGAGGCGAAAACGCTGAGGTGCCGACCTATGATTTGTGACA--CACTCAA 457
Db      835566 GCTTAAAAAATATGATGATGCTGCCAGCCCGACGTTTGTGATTTGTGTTAGTAATGAT 835507
QY      458 GGTTCACACAGACCAACGGTGTCTACCCCTGGGGACGAGTCTGTGTTGAGGACATCTTGG 517
Db      835506 CTAAACCGAGATGAGAAACGATCTATCCATTAAGTTATTTCTGTGAAAGGATTTCTGG 835447
QY      518 TGTTCACAGCCGAGAGATTCGGGACATGTTCCACTGCGCCTTTGTTGAGAACCGACT 577
Db      835446 TCTTTGAAATCAAGAACTTAGAGATCTTATGATTAATGAGATCTTTGTGACACCGATG 835387
QY      578 CCGACTCAGCGCTGTCTCGAAGAGTTCTCCCGGACCTGC--GCCGAGGAGGAGCCTGG 634
Db      835386 CTGATTAAGGATCACTACGCCGCTATGTTGAGATGTTCAAGAACAAAGGAGATAGCCTGG 835327
QY      635 AGCAGATTTTCAGCGAGTACACCACTCTGTAAGCCGGCCTTCGAGAGATTCGCTCTGC 694
Db      835326 ACTGCATCATGCTCTGTTATCTTTCTATAGTAAAGCCATATGCAATGAAATTTATAGAC 835267
QY      695 CGACAAAGAAGTATGCCGATGCGATGATCATCCACGA 729
Db      835266 CGACTCGAAATATGCTGATATCATTTGATGACATGGA 835232

RESULT 8
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: 'RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:

```

```
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 999149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9p-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.3%; Score 54.4; DB 1; Length 7218;
Matches 10; Conservative 226; Mismatches 152; Indels 0; Gaps 0;

49 GAAGGGGCGGGCGCGGACCGCATGCGCGAGCGGAGCGGAGATGCTTCGCGCGGA 108
1434 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1375
109 GCGCAACATCTGCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGATGCTTCG 168
1374 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1315
169 ATGAGGGGTGAGCGCGGCGCATGCGAGGGAAGTCGACCGTGTGTGAGAGATGATGAG 228
1314 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1255
229 TTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 288
1254 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1195
289 AGGTTCAAGAGCTCTGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
1194 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1135
349 TTGACCATCCAGATGCTTGTATATGATTTGTATGACAGAGACTGGAAGAACATGCTG 408
1134 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075
409 GAGGCGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
1074 RRRRRRRATCGCAAGTCTCCCTGACCT 1047

RESULT 9
US-08-990-140-1
Sequence 1, Application US/08990140A
Patent No. 6093795
GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Ruben, Steven M.
APPLICANT: Sonenberg, Nahum
APPLICANT: Method, Nathalie
APPLICANT: Rom, Eran
TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) and Human
FILE REFERENCE: 1488.070001
CURRENT APPLICATION NUMBER: US/08/990,140A
CURRENT FILING DATE: 1997-12-12
EARLIER APPLICATION NUMBER: US 60/033,151
```

```
EARLIER FILING DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3032
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (97)..(2718)
US-08-990-140-1

Query Match
Best Local Similarity 3.3%; Score 53.6; DB 3; Length 3032;
Matches 86; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

4 GGGTGCCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 63
175 GAGCCGCGCGACCGAGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
64 GGGACCCGATGCGCGGAGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGGAG 123
235 GGGACCGAGGCTTCAGTGAAGAGTGGGATGCGGAGCGCGCGCGCGCGGAGCGGAG 294
124 AGCCCGCGCGGAGGCGGA 143
295 AGGACCGAGCGGCGCGGA 314
```

```
RESULT 10
US-09-546-238-1
Sequence 1, Application US/09546238
Patent No. 6316225
GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Ruben, Steven M.
APPLICANT: Sonenberg, Nahum
APPLICANT: Method, Nathalie
APPLICANT: Rom, Eran
TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) Polynucleotides
FILE REFERENCE: 1488.070002
CURRENT APPLICATION NUMBER: US/09/546,238
CURRENT FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: US 60/033,151
PRIOR FILING DATE: 1996-12-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3032
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (97)..(2718)
US-09-546-238-1
```

```
Query Match
Best Local Similarity 3.3%; Score 53.6; DB 4; Length 3032;
Matches 86; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

4 GGGTGCCTTCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 63
175 GAGCCGCGCGACCGAGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 234
64 GGGACCCGATGCGCGGAGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGGAG 123
235 GGGACCGAGGCTTCAGTGAAGAGTGGGATGCGGAGCGCGCGCGCGCGGAGCGGAG 294
124 AGCCCGCGCGGAGGCGGA 143
295 AGGACCGAGCGGCGCGGA 314
```

```

RESULT 11
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
;
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB166P3
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 3.2%; Score 51.8; DB 4; Length 1830121;
Best Local Similarity 46.3%; Pred.No.0.012;
Matches 207; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY 346 AATTTCACCATCCAGATGCTTTGATTAATGATTTGATGACACGACTGGAAGAATC 405
| | | | |
DB 146192 AATTGAGATCATCCAAACCTGATGATCGGATTTACTTATCCAACTTTAAAAAATCTA 146251
| | | | |
QY 406 GTGGAGGGGAAAAAGGTGAGGTGCGGACCTATGATTTGTGTGACACACTCAAGTTACCA 465
| | | | |
DB 146252 AAAAAATGGAGTGCAGTAGATGTGCTGTTTATAGCTATGTAAAGAACTCCGACCAAC 146311
| | | | |
QY 466 GAGACCAAGGTGTCACCTCGGAGACGTGTCTGTTGAGGGGCACTTGTTGTTTAC 525
| | | | |
DB 146312 GAAACACCACTTTTACACACAAAAAGCAATGTAATTTTAAAGGAAATTTATTACTTACC 146371
| | | | |
QY 526 AGCCAGAGATCCGGGACATGTTCCACTCGCGCTTTCGTGAGACACCGACTCCGAGTC 585
| | | | |
DB 146372 GATGAACGAGTACGCAATTAAGCCGATATTTCGTATTGTTGTATACACACACACTGTAATTT 146431
| | | | |
QY 586 AGGCTGTCTCGAAGAATTTCCGGGACGTGCGC---CGAGGGAGGGGACTGAGACAGATT 642
| | | | |
DB 146432 TGTTCATCTCCGCGTTTACCAACGTGATATGAGAAACGTGTGCTCTTACATCAGTG 146491
| | | | |
QY 643 CTGACGACGATACACACCTTCGTGGAAGCGGCGCTTGAGAGAGTTCTGCGCCGACAAAG 702
| | | | |
DB 146492 ATTGATCAATATCTGCGCAACCGGCGCTCCAGTGTCTTACAAATTTATTGAGCCGCTTAAA 146551

```

```

QY      703 AAGTATGCGAATGTATCATCTCCACAGAGAGTGAGCAATATGTTGCATCAACCTGATC 762
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      146552 CAATTTCGGAGTAATTTGCTACTTCTCTGTGGTGAATAAAATCGCATTCGATCATATGTTA 1466111
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      763 GTGCAGCACATCCAGACATTTCTGAT 789
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      146612 AAAGCTCAAATCTTTCATTATTGAAAT 146638
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| |

RESULT 12
US-09-643-990A-1
; Sequence 1, Application US//09643990A
; Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
          Mark D. Adams
          Owen White
          Hamilton O. Smith
          J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
                    the Haemophilus influenzae Rd Genome, Fragments
                    Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match           3.2%; Score 51.8; DB 4; Length 1830121;
Best Local Similarity 46.3%; Pred. No. 0.012;
Matches 207; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

QY      346 AATTGTGACCATCCGAGATGCCCTTTATATATGTTTGATGATGACACACTCAAGGTTAACCA 405
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      146192 AATTAGCATCATCCAACTCTATGATGCGCGATTACTTATCCAACTTTAAAAAATCTA 1462511
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      406 GTGAGGAGCAAAAAGCGTGAGGTGCGACACTATGATTTTGTGACACACTCAAGGTTAACCA 465
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      146252 AAAAATGGCAGTGCAGTAGATGTGCTCTTTTATAGCTATGTGAAACATATCTGCGACCAAC 14631111

```


NAME: Kempier, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 4092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 200..3028
OTHER INFORMATION:
NAME/KEY: Human ROR2
LOCATION: 1...4092
OTHER INFORMATION:
US-08-469-537A-106

Query Match 3.1%; Score 50.4; DB 2; Length 4092;
Best Local Similarity 48.9%; Pred. No. 0.0015;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 222 CATGAGGTTGCTGGAGACAGACGAGTGAACAGCGGAGCGGAGGAGTGCATCTGAG 281
DB 1552 CATGGAATGCCCTCATTTACACGACAAACAGCCAACTCAAAGATCAGCTGTC 1611
QY 282 CCAGACAGGTTCTACAAAGTCTGACGCGACAGACAGAGCCAAAGCCTTGAAGACA 341
DB 1512 TGCCGTGAGGTTCAATGAGGAGGAGTGGAGAGACCGGTTTGGAAGTCTACAAAGTCA 1671
QY 342 GTACAATTTTACCATCCAGATGCTTTGATAATGATTGTGACAGAGATCTGGAAGAA 401
DB 1672 CTTGTCGGCCCTGCCCCGGGGAGACAGACCCAGGCTGTGCCATCAAAAGCTGAAGGA 1731
QY 402 CATGTTGAGGGCAAAAACGTTGAGAGTCCGACCTATGATTTTGAACACATCAAGGTT 461
DB 1732 CAAAGCGAGGGGCCCCCTCCGGAGAGAGTCCCGCATGAGCTATGCTGCGAGACGGCT 1791
QY 462 ACCAGAGACACAGGTGTCTACCTGCGAGAGTGT 497
DB 1792 GCAACACCCCAACGTCGTCTGCTGCGGCGTGT 1827

RESULT 15
US-07-642-734C-3
Sequence 3, Application US/07642734C
Patent No. 5824513
GENERAL INFORMATION:
APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: Mcalpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckerts, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20235 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238
FEATURE:
NAME/KEY: CDS
LOCATION: 19..10722
OTHER INFORMATION: /codon_start= 19
OTHER INFORMATION: /function= "gene eryA"
OTHER INFORMATION: /product= "eryA ORF2 encoding modules 3 & 4 for
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..4470
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 97..1482
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1693..2670
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3406..3921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4171..4428
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..10722
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..5847
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl:Acyl synthase domain of module"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6054..7026
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7165..9216
OTHER INFORMATION: /function= "approximate span of

```

1 OTHER INFORMATION: dehydratase and enoylreductase domains
2 FEATURE:
3 NAME/KEY: misc_feature
4 LOCATION: 9433..9984
5 OTHER INFORMATION: /function= "approximate span
6 OTHER INFORMATION: beta-ketoreductase of module 4"
7 FEATURE:
8 NAME/KEY: misc_feature
9 LOCATION: 10225..10483
10 OTHER INFORMATION: /function= "approximate span of
11 OTHER INFORMATION: acyl carrier domain of module 4"
12 FEATURE:
13 NAME/KEY: CDS
14 LOCATION: 10723..20235
15 OTHER INFORMATION: /codon_start= 10723
16 OTHER INFORMATION: /function= "gene =eryA"
17 OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
18 OTHER INFORMATION: 6-deoxyerythronolide B formatio"
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: 10723..15165
22 OTHER INFORMATION: /function= "approximate span of
23 OTHER INFORMATION: module 5"
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: 10831..12174
27 OTHER INFORMATION: /function= "approximate span of
28 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
29 FEATURE:
30 NAME/KEY: misc_feature
31 LOCATION: 12379..13350
32 OTHER INFORMATION: /function= "approximatr span of
33 OTHER INFORMATION: acyltransferase domain of module 5"
34 FEATURE:
35 NAME/KEY: misc_feature
36 LOCATION: 14062..14610
37 OTHER INFORMATION: /function= "approximate span of
38 OTHER INFORMATION: beta-ketoreductase of module 5"
39 FEATURE:
40 NAME/KEY: misc_feature
41 LOCATION: 14857..15114
42 OTHER INFORMATION: /function= "approximate span of
43 OTHER INFORMATION: acyl carrier domain of module 5"
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: 15166..20235
47 OTHER INFORMATION: /function= "approximate span of
48 OTHER INFORMATION: module 6"
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: 15172..16569
52 OTHER INFORMATION: /function= "approximate span of
53 OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
54 FEATURE:
55 NAME/KEY: misc_feature
56 LOCATION: 16768..17721
57 OTHER INFORMATION: /function= "approximate span of
58 OTHER INFORMATION: acyltransferase domain of module 6"
59 FEATURE:
60 NAME/KEY: misc_feature
61 LOCATION: 18379..18921
62 OTHER INFORMATION: /function= "approximate span of
63 OTHER INFORMATION: beta-ketoreductase domain of module 6"
64 FEATURE:
65 NAME/KEY: misc_feature
66 LOCATION: 19149..19398
67 OTHER INFORMATION: /function= "approximate span of
68 OTHER INFORMATION: acyl carrier domain of module 6"
69 FEATURE:
70 NAME/KEY: misc_feature
71 LOCATION: 19492..20235
72 OTHER INFORMATION: /function= "approximate span of
73 OTHER INFORMATION: thioesterase domain of module 6"

```

Query Match	2.9%	Score 46.6	DB 1	Length 20235
Best Local Similarity	47.2%	Pred. No. 0.034		
Matches 142	Conservative 0	Mismatches 159	Indels 0	Gaps 0
QY	33	GCGCGCCGGGCGCGGGAGAGGCGGGCGCGGGACCCGATGCCGGGAGCGGAGGCCGA	92	
Db	5367	GGGCGTCGGGCGCGGACGTCGACGTCGTCGAGAGCCACCGGACCGGGCACCGAGCTCGG	5426	
QY	93	GATGGCTTCGGCGGGAGGCGGAAGCTCGAGAGCCCCCGCGCGGAGCCGACCTTCGCA	152	
Db	5427	CGACCCGATGAGGCGCGGCGCTCATTCGCGGACCTTAGCGCCAGGACCGCGACCGGCT	5486	
QY	153	CCAGCGGCCCTCTCTGATAGGGGAGAGCGGCGGCGGACCTGCCAGCGGGAAGTGAACGTCG	212	
Db	5487	GCGGCTCGGCTCGGTGGAAGCAACATCGGCGCACCGAGGCCGCGGCGCGCGCGG	5546	
QY	213	TGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACAGGGGACGCGGAAGTGT	272	
Db	5547	CGTATCAAGGTCGTGCTGGGATGCGGACAGGATCTGCCCCGCTGTTCACGCCGA	5606	
QY	273	CATGCTGAGCCAGGACGTTCTACCAAGGTCCTGACGGCAGAGCAGAGGCCAGGCTT	332	
Db	5607	CGAGCTGTCCCGCAGATCGACTGGGAGTCGGGGGCGGTGAGGTGCTGCGCGAGAGGT	5666	
QY	333	G 333		
Db	5667	G 5667		

Search completed: November 25, 2003, 02:03:52
Job time : 125.319 secs

Tr: -

This Page Blank (uspto)

QY	18	GGCGCACTGCGACGGGAAAGTCACCGTGTGTGAAAGATCATGGAATTTCTGGGACAG	240
Dp	181	GGCGGCATTCGCCACGGGAAAGTCACCGTGTGTGAAAGATCATGGAATTTCTGGGACAG	240
QY	241	AACGAGGTGAAACAGCGGCACGCGAAGGTGTCTATCTTGAGCCAGGACAGGTTCTAACG	300
Dp	241	AACGAGGTGAAACAGCGGCACGCGAAGGTGTCTATCTTGAGCCAGGACAGGTTCTAACG	300
QY	301	GTCTCGAAGGAGAGCACAAGGCCCTTGAAAGGACAGTACAAATTTTGAACATCA	360
Dp	301	GTCTCGAAGGAGAGCACAAGGCCCTTGAAAGGACAGTACAAATTTTGAACATCA	360
QY	361	GATGCGCTTGAATATGATTTTGATGACACAGACTCTGAAGAAACATCGTGAAGGCAAAACG	420
Dp	361	GATGCGCTTGAATATGATTTTGATGACAGACTCTGAAGAAACATCGTGAAGGCAAAACG	420
QY	421	GTGGAAGTCCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCGTGTCTC	480
Dp	421	GTGGAAGTCCCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCGTGTCTC	480
QY	481	TACCCGTGCGAGCGTGTCTGTGTTGAGGACACTCTGATGTTCTACAGCAGAGAGATCCGG	540
Dp	481	TACCCGTGCGAGCGTGTCTGTGTTGAGGACACTCTGATGTTCTACAGCAGAGAGATCCGG	540
QY	541	GACATGTTCCACCTGCGGCTCTTCGTGTGACACCGACTCCGACGTCAAGCTGTCTGAAGA	600
Dp	541	GACATGTTCCACCTGCGGCTCTTCGTGTGACACCGACTCCGACGTCAAGCTGTCTGAAGA	600
QY	601	GTTCTCCGCGGACGTGCGCGCGAGGAGGACCTTGAGAGCAATTTGACGCAATACCAACC	660
Dp	601	GTTCTCCGCGGACGTGCGCGCGAGGAGGACCTTGAGAGCAATTTGACGCAATACCAACC	660
QY	661	TTCTGTGAAGCCGGGCTTCGAGAGAGTTCGTGCTGCGACGACAAAGAGTATGCCGATGTATC	720
Dp	661	TTCTGTGAAGCCGGGCTTCGAGAGAGTTCGTGCTGCGACGACAAAGAGTATGCCGATGTATC	720
QY	721	ATCCCAAGAGAGGTGAGCAATATGTTGTGTCATCAACTTGATCTGTGACAGACATCCAGGAC	780
Dp	721	ATCCCAAGAGAGGTGAGCAATATGTTGTGTCATCAACTTGATCTGTGACAGACATCCAGGAC	780
QY	781	ATTCTGAATGTTGATCTGCAAAATGAGACCGAGAGAGGTCCAAATGGGCGGAGCTACAG	840
Dp	781	ATTCTGAATGTTGATCTGCAAAATGAGACCGAGAGAGGTCCAAATGGGCGGAGCTACAG	840
QY	841	CGGACCTTTCTGAGCCAGGAGGACCACTTGAGATGCTGACCTCTGGCAAAACGCTCATAT	900
Dp	841	CGGACCTTTCTGAGCCAGGAGGACCACTTGAGATGCTGACCTCTGGCAAAACGCTCATAT	900
QY	901	TTGAGATCTCAGACGACACCCCACTGAAGAGGCTGCGAGCCCTCAAGGGACAGTCCCGGCC	960
Dp	901	TTGAGATCTCAGACGACACCCCACTGAAGAGGCTGCGAGCCCTCAAGGGACAGTCCCGGCC	960
QY	961	CGGACATGTGTGTTCAAGGACCTGAGCCTTGAGGACGCGCCACCCACACTGCTCTCTCTC	1020
Dp	961	CGGACATGTGTGTTCAAGGACCTGAGCCTTGAGGACGCGCCACCCACACTGCTCTCTCTC	1020
QY	1021	GGCGGACCCCAAGGGAGGTGTAGCCAGCGAGGCTTCTCACTCAAGAGTGAATCTCAGA	1080
Dp	1021	GGCGGACCCCAAGGGAGGTGTAGCCAGCGAGGCTTCTCACTCAAGAGTGAATCTCAGA	1080
QY	1081	TGTGTCACTCAGACTCAACTTGTCTGAGGACACTGACAGGCGTTCCTGAAGTTTTCAAGCCAC	1140
Dp	1081	TGTGTCACTCAGACTCAACTTGTCTGAGGACACTGACAGGCGTTCCTGAAGTTTTCAAGCCAC	1140
QY	1141	TTAGGCTGTGCGGTTTTAAAGATCCCTAGGTCACTGAAGAAATGCCACAGAAATGTGCA	1200
Dp	1141	TTAGGCTGTGCGGTTTTAAAGATCCCTAGGTCACTGAAGAAATGCCACAGAAATGTGCA	1200
QY	1201	GGAAGCCTGGGAGGCTTCTGTGAGGAATGTAGGCACTTATTTGGGGAAATTTGAGAGAC	1260
Dp	1201	GGAAGCCTGGGAGGCTTCTGTGAGGAATGTAGGCACTTATTTGGGGAAATTTGAGAGAC	1260

Qy	1261	AGCCTTAACACTGGCTGGCCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGATT	1320
Db	1261	AGCCTTAACACTGGCTGGCCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGATT	1320
Qy	1321	TTTTCCAGTCTGATCTGTTGTTCTTACACACTCACACATTAATCTCAAAAGTTTGTGAA	1380
Db	1321	TTTTCCAGTCTGATCTGTTGTTCTTACACACTCACACATTAATCTCAAAAGTTTGTGAA	1380
Qy	1381	AGTACTTTCCTTTTTCATGTATACATGTCCTCATGTTTCTGTTTCTGTTCATTAACA	1440
Db	1381	AGTACTTTCCTTTTTCATGTATACATGTCCTCATGTTTCTGTTTCTGTTCATTAACA	1440
Qy	1441	CAAAGCTGTGTGTGGCCTTACAAACCTTAATTCATGACCCAGTGGTTGCACTCCAGCGTG	1500
Db	1441	CAAAGCTGTGTGTGGCCTTACAAACCTTAATTCATGACCCAGTGGTTGCACTCCAGCGTG	1500
Qy	1501	GCCCTACACGGATATGGGGAGCACCTGAGGGATGTTTTCCCCCTTGCTGTGTGCTTTAAAG	1560
Db	1501	GCCCTACACGGATATGGGGAGCACCTGAGGGATGTTTTCCCCCTTGCTGTGTGCTTTAAAG	1560
Qy	1561	GCAAGAAAGGAGGCGGATGCGCTCTGGAAGCACCCAGATACACCCAGGCTTGTGCGGGG	1620
Db	1561	GCAAGAAAGGAGGCGGATGCGCTCTGGAAGCACCCAGATACACCCAGGCTTGTGCGGGG	1620
Qy	1621	CCAG 1624	
Db	1621	CCAG 1624	

RESULT 2
 US-09-833-381-2048
 : Sequence 2048, Application US/09833381
 : Patent No. US20020132090A1
 : GENERAL INFORMATION:
 : APPLICANT: ROBINSON, Keith E.
 : TITLE OF INVENTION: NO. US20020132090A1 Nucleic Acid and Protein Homologs
 : FILE REFERENCE: 5800-119
 : CURRENT APPLICATION NUMBER: US/09/833,381
 : CURRENT FILING DATE: 2001-04-11
 : PRIOR APPLICATION NUMBER: 09/516, 448
 : PRIOR FILING DATE: 2000-02-29
 : NUMBER OF SEQ. ID NOS: 2050
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 2048
 : LENGTH: 1648
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc_feature
 : LOCATION: (1)...(1648)
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-833-381-2048

Good Data

[illegible]

QY 254 AGCGGACGGGAAAGTGTGATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGGCAG 313
 DB 244 AGGCGACGGGAAAGTGTGATCTGAGCCAGGACAGGTTCTCAAGAGTCTGAGCCAG 303
 QY 314 AGGAGAGGCGCAAGGCGCTTGAAGAGACATATTTTGAACATTCAGATGCTTTGATA 373
 DB 304 AGCAGAAAGGCGCAAGGCGCTTGAAGAGACATATTTTGAACATTCAGATGCTTTGATA 363
 QY 374 ATGATTTGATGACACAGACTCTGAAGAACATCTGTGAGGGGCAAAACGCTGAGAGTCCGA 433
 DB 364 ATGATTTGATGACACAGACTCTGAAGAACATCTGTGAGGGGCAAAACGCTGAGAGTCCGA 423
 QY 434 CCTATGATTTTGTGACACACTCAAGTTACAGAGACCAAGGAGTGTACCTGCGGAGC 493
 DB 424 CCTATGATTTTGTGACACACTCAAGTTACAGAGACCAAGGAGTGTACCTGCGGAGC 483
 QY 494 TGGTTCGTGTTGAGGGGATCTGTGTGTTCTACAGCCAGAGATCCGGACATGTTCCACC 553
 DB 484 TGGTTCGTGTTGAGGGGATCTGTGTGTTCTACAGCCAGAGATCCGGACATGTTCCACC 543
 QY 554 TGGGCGCTCTTCTGAGACACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTCTCCGGAGC 613
 DB 544 TGGGCGCTCTTCTGAGACACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTCTCCGGAGC 603
 QY 614 TGGCGCCAGGAGGAGGACCTGAGAGATCTGACGAGTACACACACTTGTGAGAGCGG 673
 DB 604 TGGCGCCAGGAGGAGGACCTGAGAGATCTGACGAGTACACACACTTGTGAGAGCGG 663
 QY 674 CTTTCGAGAGATTCTGCTGCGCCGACAAAGAAAGTA-TGCCAGATGTATCATCCACGAGGA 732
 DB 664 CTTTCGAGAGATTCTGCTGCGCCGACAAAGAAAGTATTCGATGTATCATCCACGAGGA 723
 QY 733 GTGAGACAAATGTGTTGCTGATCAACCTGATGTGAGACACATCCAGACATTTGATGAT 792
 DB 724 GTGAGACAAATGTGTTGCTGATCAACCTGATGTGAGACACATTTGATGATGAT 783
 QY 793 GACATCTGCAAAATGGGACCGAGAGGGGTCCAAATGGGGGAGGCTACAAAGGGAGCCTTTCT 852
 DB 784 GACATCTGCAAAATGGGACCGAGAGGGGTCCAAATGGGGGAGGCTACAAAGGGAGCCTTTCT 843
 QY 853 GAGCCAGGGGAGGACCCCTGGAGTGTGACCTCTGCAAAACGCTCACATTTGAGTCCAGC 912
 DB 844 GAGCCAGGGGAGGACCCCTGGAGTGTGACCTCTGCAAAACGCTCACATTTGAGTCCAGC 903
 QY 913 AGCAGACCCCACTGAGAGGGCTGCCAGGCTTCAAGGGAGGCTTCCGCCGCGACATGTGT 972
 DB 904 AGCAGACCCCACTGAGAGGGCTGCCAGGCTTCAAGGGAGGCTTCCGCCGCGACATGTGT 963
 QY 973 TCAGGGAGTGAAGCTGGGGAGCGCCACACCCACTGCTTCTCTCGGCGGACCCGAG 1032
 DB 964 TCAGGGAGTGAAGCTGGGGAGCGCCACACCCACTGCTTCTCTCGGCGGACCCGAG 1023
 QY 1033 GGGAGTGTGAGACGAGGCTTCTCTCACTCAGAGAGTGAACCTCAGATGTGCACTCAG 1092
 DB 1024 GGGAGTGTGAGACGAGGCTTCTCTCACTCAGAGAGTGAACCTCAGATGTGCACTCAG 1083
 QY 1093 ACTCAACTTCTGAGACACTGACAGGCGTTCCTGAGGTTTTCAGCCACTTGAAGCTGTTG 1152
 DB 1084 ACTCAACTTCTGAGACACTGACAGGCGTTCCTGAGGTTTTCAGCCACTTGAAGCTGTTG 1143
 QY 1153 CGGTTTAAAGATCCCTTAGGTCACTGAGAGAAATGSCACAGAAATGTGAGAGAAACCTGGGA 1212
 DB 1144 CGGTTTAAAGATCCCTTAGGTCACTGAGAGAAATGSCACAGAAATGTGAGAGAAACCTGGGA 1203
 QY 1213 GGGCTTCTGAGAGATGTGAGGACATTAATTTGGGAAATTTGAGAGAGACAGCTTGAACACT 1272
 DB 1204 GGGCTTCTGAGAGATGTGAGGACATTAATTTGGGAAATTTGAGAGAGACAGCTTGAACACT 1263
 QY 1273 GGGCTGAGCTATTTTGTGAGACATGTAACCCACAGTGGGAGAGATTTTTCAGAGCTG 1332
 DB 1264 GGGCTGAGCTATTTTGTGAGACATGTAACCCACAGTGGGAGAGATTTTTCAGAGCTG 1323
 QY 1333 ATCTGTTCTTACACACTCAACATTACTCAAAAGTTTGTGGAACAAGTACTTTCCTT 1392

DB 1324 ATCTGTTCTTACACACTCAACACATTAAGTATTTGTGAACAAGTACTTTCCTT 1383
 QY 1393 TTTTACATGTTACATGATCTCTCATGTTTCTGTTTCTGTTTCAACACAGGCTGTTG 1452
 DB 1384 TTTTACATGTTACATGATCTCTCATGTTTCTGTTTCTGTTTCAACACAGGCTGTTG 1443
 QY 1453 TGGCCCTAACAACCTAATTTTCAATGACCCAGTGTTCATTCAGAGTGTGCTTACAGAGAT 1512
 DB 1444 TGGCCCTAACAACCTAATTTTCAATGACCCAGTGTTCATTCAGAGTGTGCTTACAGAGAT 1503
 QY 1513 ATGGGAGGACACTGAGAGATGTTTTCCTGCTTGTGCTTAAAGGACAGAGAGCA 1572
 DB 1504 ATGGGAGGACACTGAGAGATGTTTTCCTGCTTGTGCTTAAAGGACAGAGAGCA 1563
 QY 1573 GGGGATGCTCTTGAAGACCCAGACATCAACCCAGGCTTGTGCGGGGCGAG 1624
 DB 1564 GGGGATGCTCTTGAAGACCCAGACATCAACCCAGGCTTGTGCGGGGCGAG 1615

RESULT 3
 US-10-037-270-546
 ; Sequence 546, Application US/10037270
 ; Publication No. US20030104529A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yang, Y. Tom
 ; APPLICANT: Yang, Y. Tom
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunding
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Tillinghast, John
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 ; FILE REFERENCE: Polypeptides
 ; CURRENT APPLICATION NUMBER: US/10/037,270
 ; PRIOR FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; NUMBER OF SEQ ID NOS: 1104
 ; SOFTWARE: pc FL_genes Version 1.0
 ; SEQ ID NO 546
 ; LENGTH: 1288
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (95)..
 ; US-10-037-270-546

Query Match 60.3%; Score 978.6; DB 14; Length 1288;
 Best Local Similarity 99.6%; Pred. No. 3,4e-285;
 Matches 981; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 29 GGGGGGCGCGCGGCGCGGCGGGAAGGGGGGCGCGGGGACCCGATGCGCGGGAGCGGAGG 88
 DB 30 GGGTCCGCGCGCGCGCGCGGGAAGGGGGGCGCGGGGACCCGATGCGCGGGAGCGGAGG 89
 QY 89 CCGAGATGCTTCCGCGGAGGCGGAAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 148
 DB 90 CCGAGATGCTTCCGCGGAGGCGGAAGCTGCGAGAGCCCGCGCGGAGCGGACCGTC 149

Good data

149 CGCACAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCGACCTGCCAGCGGGAATGCGACCG 208
150 CGCACAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCGACCTGCCAGCGGGAATGCGACCG 209
209 TGTGTGAGGAAGATCATGAGTGTGCTGGGACAGAAAGAGTGAAGAGCGGCGGAGG 268
210 TGTGTGAGGAAGATCATGAGTGTGCTGGGACAGAAAGAGTGAAGAGCGGCGGAGG 269
269 TGTGTATCTTGGAGCCAGAGCAGGTTCTTCAAGAGTCTTGAACGGCAGACAGAAAGGCGCAAGG 328
270 TGTGTATCTTGGAGCCAGAGCAGGTTCTTCAAGAGTCTTGAACGGCAGACAGAAAGGCGCAAGG 329
329 CTTTGAAGAAGCAGTATCAATTTTGAACATCCAGATGCTTGAATGAATGATTTGATGACA 388
330 CTTTGAAGAAGCAGTATCAATTTTGAACATCCAGATGCTTGAATGAATGATTTGATGACA 389
389 GGAATCTGAAGAAGATCATGAGTGTGCTGGGACAGAAAGAGTGAAGAGCGGCGGAGTGA 448
390 GGAATCTGAAGAAGATCATGAGTGTGCTGGGACAGAAAGAGTGAAGAGCGGCGGAGTGA 449
449 CACACTCAAGGTTACAGAGACCAAGGTGTCTACCTTGGGACGTGGTCTGTTGAGG 508
450 CACACTCAAGGTTACAGAGACCAAGGTGTCTACCTTGGGACGTGGTCTGTTGAGG 509
509 GCATCTTGTGTGTTTGAAGCCAGAGAGATCCGGGACATGTTCCAGTGGCGCTCTTCTGTTG 568
510 GCATCTTGTGTGTTTGAAGCCAGAGAGATCCGGGACATGTTCCAGTGGCGCTCTTCTGTTG 569
569 ACACCGAGCTCCGAGGTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGGCGGAGGAGG 628
570 ACACCGAGCTCCGAGGTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGGCGGAGGAGG 629
629 ACCTGAGCAGATTTCTGACGAGTACACCACTTCTGTAAGCCGGCTTGGAGAGTTCT 688
630 ACCTGAGCAGATTTCTGACGAGTACACCACTTCTGTAAGCCGGCTTGGAGAGTTCT 689
689 GCGTCGCGGACAAAGAGATGCGGATGATCATCCACGAGAGTGAACAAATGATGTTG 748
690 GCGTCGCGGACAAAGAGATGCGGATGATCATCCACGAGAGTGAACAAATGATGTTG 749
749 CCATCAACCTGATGTCGAGCAGATCCAGAGATTTCTGAATGATGATCTGCAATGAGC 808
750 CCATCAACCTGATGTCGAGCAGATCCAGAGATTTCTGAATGATGATCTGCAATGAGC 809
809 ACCGAGAGAGGTCCTAATGAGCGAGTCAACAGCGACCTTTTCTGAGCCAGGAGCAACC 868
810 ACCGAGAGAGGTCCTAATGAGCGAGTCAACAGCGACCTTTTCTGAGCCAGGAGCAACC 869
869 CTGGGATGCTGACCTCTGGGACAAAGGTCAATTTGAGTCCAGAGAGAGCCCACTGAG 928
870 CTGGGATGCTGACCTCTGGGACAAAGGTCAATTTGAGTCCAGAGAGAGCCCACTGAG 929
929 GGGGTGCGGAGCTCTCAGGAGAGTCTCCCGCCGAGATGTTGTTTCAAGGACTGAGCTG 988
930 GGGGTGCGGAGCTCTCAGGAGAGTCTCCCGCCGAGATGTTGTTTCAAGGACTGAGCTG 989
989 GGGAGCGCCACCAACACCTGCT 1013
990 GGGAGCGCCACCAACACCTGCT 1014

RESULT 4
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764, 877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804

Query Match 54.2%; Score 880; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 7.1e-255;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

745 GTTGCCATCAACCTGATCGTGCAGCAGATCCAGAGATTTCTGAATGATGATCTGCAAA 804
3901 GTTGCCATCAACCTGATCGTGCAGCAGATCCAGAGATTTCTGAATGATGATCTGCAAA 3960
805 TGGCACCGAGAGAGGTCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGAGC 864
3961 TGGCACCGAGAGAGGTCTCAATGGCGGAGCTCAAGCGGACCTTTTCTGAGCCAGGGAGC 4020
865 CACCTGGGATGCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGCAGACCCAC 924
4021 CACCTGGGATGCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGCAGACCCAC 4080
925 TGAGGGGCTGCCGAGCTCAGGGCAGGTCTCCCGCCGAGATGTTGTTCAAGGACTGAG 984
4081 TGAGGGGCTGCCGAGCTCAGGGCAGGTCTCCCGCCGAGATGTTGTTCAAGGACTGAG 4140
985 CTTGGGAGCGCCACCCACCAACCCACTGCTTCTTCTGGGCAACCCAGGGAGGTGTTAGC 1044
4141 CTTGGGAGCGCCACCCACCAACCCACTGCTTCTTCTGGGCAACCCAGGGAGGTGTTAGC 4200
1045 AGGAGAGCTTCTCTCACTCAGAGTGAAGAACTCAATGTTGTCATCTAGCTCAACTTGTCT 1104
4201 AGGAGAGCTTCTCTCACTCAGAGTGAAGAACTCAATGTTGTCATCTAGCTCAACTTGTCT 4260
1105 GGGACACTGACAGAGCTTCTCTGAGTTCAGCACTTGAAGCTGTTGGGTTTAAAGAT 1164
4261 GGGACACTGACAGAGCTTCTCTGAGTTCAGCACTTGAAGCTGTTGGGTTTAAAGAT 4320
1165 CCCTTAGGTCACCTGAAATGCGACAGATGTCAGAAAGCTTGGAGGCTTCTGTGAG 1224
4321 CCCTTAGGTCACCTGAAATGCGACAGATGTCAGAAAGCTTGGAGGCTTCTGTGAG 4380
1225 GAATGTAGGACATTAATGGGGAATTTGAGAGAGACGCTTGAACACTGCTGCTGAT 1284
4381 GAATGTAGGACATTAATGGGGAATTTGAGAGAGACGCTTGAACACTGCTGCTGAT 4440
1285 GTTTGTGACAGTGAACCAAGTGGAGAGATTTTTCAGTCTGATCTGATCTTCA 1344
4441 GTTTGTGACAGTGAACCAAGTGGAGAGATTTTTCAGTCTGATCTGATCTTCA 4500
1345 CACACTCAGACATTAATCAAAAGTTTGTGAACAGTACTTCTTTTAAATGTTA 1404
4501 CACACTCAGACATTAATCAAAAGTTTGTGAACAGTACTTCTTTTAAATGTTA 4560
1405 CATGTCTCATGTTTCTGTTTCTGTTTCTAATCAACAAAGCTGTTGCTGCTCAAC 1464
4561 CATGTCTCATGTTTCTGTTTCTGTTTCTAATCAACAAAGCTGTTGCTGCTCAAC 4620
1465 CTAAATTCATGACCCAGTGTGTCAGTCCAGGTCCTTACACGATATAGGAGAGCCAC 1524
4621 CTAAATTCATGACCCAGTGTGTCAGTCCAGGTCCTTACACGATATAGGAGAGCCAC 4680
1525 TGAGGAGTGTTCCTCCCTGTTGCTTGAAGAGCAGAAAGCAGAGCGGATGCTT 1584
4681 TGAGGAGTGTTCCTCCCTGTTGCTTGAAGAGCAGAAAGCAGAGCGGATGCTT 4740
1585 GGAAGCACCAGCATCAACCCAGGCTTGTGGGGGCCAG 1624
4741 GGAAGCACCAGCATCAACCCAGGCTTGTGGGGGCCAG 4780


```
Db 541 GAGCAGATTGAGCAGCAACCACTTCGTGAAGCCGGCTTGAGAGTTCTGCTG 600
Qy 694 CCGCAAGAAGTATCCCGATGTGATCATCCACAGAGATGAGCAATATGTTGCCATC 753
Db 601 CCGCAAGAAGTATGCGATGTGATCATCCACAGAGATGAGCAATATGTTGCCATC 660
Qy 754 AACGTATGTCAGCAGCATCCAGACATTTCTGAATGTTGATCATCTGCAATGAGCAGCA 813
Db 661 AACGTATGTCAGCAGCATCCAGACATTTCTGAATGTTGATCATCTGCAATGAGCAGCA 720
Qy 814 GAGAGTCCAAATGAGCGAGCTCAAGCGGACCTTTTCTGAGCGAGGAGCACCCTGGG 873
Db 721 GAGAGTCCAAATGAGCGAGCTCAAGCGGACCTTTTCTGAGCGAGGAGCACCCTGGG 780
Qy 874 ATGTCAGCTTCGCAAAACGTCACATTTGAGATCCAGACAGACCCCACTGA 927
Db 781 ATGTCAGCTTCGCAAAACGTCACATTTGAGATCCAGACAGACCCCACTGA 834
```

RESULT 7

```
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9,1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g11899252, EVALU0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALU0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU2.00e-25
US-10-029-386-2183
```

Query Match 29.9%; Score 485; DB 12; Length 510;
Best Local Similarity 100.0%; Pred. No. 5.1e-136;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 745 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGTTGACATTCGAAA 804
Db 26 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGTTGACATTCGAAA 85
Qy 805 TGGCAGCAGAGAGGCTCCATATGGGCGAGCTCAAGAGGAGCCTTTCTGAGCGAGGAGC 864
Db 86 TGGCAGCAGAGAGGCTCCATATGGGCGAGCTCAAGAGGAGCCTTTCTGAGCGAGGAGC 145
Qy 865 CACCTGAGATGCTGACCTTGCGAAACGATCACATTTGAGTCCAGACAGACCCAC 924
Db 146 CACCTGAGATGCTGACCTTGCGAAACGATCACATTTGAGTCCAGACAGACCCAC 205
Qy 925 TGAAGGCTGCGAGGCTCAAGGCAAGTCTCCCGCCGCGCATGTGTCTTCAGGAGCTGAG 984
Db 206 TGAAGGCTGCGAGGCTCAAGGCAAGTCTCCCGCCGCGCATGTGTCTTCAGGAGCTGAG 265
```

```
Qy 985 CTTGGGAGGCCACCCACACCACTTCTCTCGGCGCACCCAGGGAGGTTAGC 1044
Db 266 CTTGGGAGGCCACCCACACCACTTCTCTCGGCGCACCCAGGGAGGTTAGC 325
Qy 1045 AGGAGGCTTCTCTCACTCAAGAGTGAATCTCAGATGTGTCACTGACCTCAACTTCT 1104
Db 326 AGGAGGCTTCTCTCACTCAAGAGTGAATCTCAGATGTGTCACTGACCTCAACTTCT 385
Qy 1105 GGGACACTGACAGGCTTCTCTGAGGTTTTCAGCCACTTAGGCTGTTTGAAGAT 1164
Db 386 GGGACACTGACAGGCTTCTCTGAGGTTTTCAGCCACTTAGGCTGTTTGAAGAT 445
Qy 1165 CCGCTGATGATCAGAGAAATGCGACAGAAATGTGACAGAAAGCTTGAGAGCTTCTGAG 1224
Db 446 CCGCTGATGATCAGAGAAATGCGACAGAAATGTGACAGAAAGCTTGAGAGCTTCTGAG 505
Qy 1225 GAATG 1229
Db 506 GAATG 510
```

RESULT 8

```
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

Query Match 26.7%; Score 432.8; DB 11; Length 472;
Best Local Similarity 98.2%; Pred. No. 3.1e-120;
Matches 437; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy 177 GAGCGCGGACACTGCGAGCGGGAAGTCGACCCGTGTGAGAAATATATGAGTTGCGG 236
Db 28 GTGCGTGGAAAGCGCAGCGGGAAGTCGACCCGTGTGAGAAATATATGAGTTGCGG 87
Qy 237 ACAGAACGAGGTGGAACAGCGGAGCGGAAGGTGATCTCTGAGCAGAGAGTTCTA 296
Db 88 ACAGAACGAGGTGGAACAGCGGAGCGGAAGGTGATCTCTGAGCAGAGAGTTCTA 147
Qy 297 CAAGGTCTGACGCGCAGAGCAAGGCAAGGCTTTGAAAGCAGTAACTATTTTGACA 356
Db 148 CAAGGTCTGACGCGCAGAGCAAGGCAAGGCTTTGAAAGCAGTAACTATTTTGACA 207
Qy 357 TCCAGATGCTTTGATTAATGATTTGATGACAGAGATCTTGAAAGAACTATGAGAGGCAA 416
Db 208 TCCAGATGCTTTGATTAATGATTTGATGACAGAGATCTTGAAAGAACTATGAGAGGCAA 267
Qy 417 AACGGTGAAGTCCGACCTATGATTTTGTGACACTCAAGGTTACAGAGACACGGT 476
Db 268 AACGGTGAAGTCCGACCTATGATTTTGTGACACTCAAGGTTACAGAGACACGGT 327
Qy 477 GGTCTACCTGCGAGAGTGGTTCTGTTTGAAGGCACTTGTGTTTCTAAGCAGAGAT 536
```

QY	Db
537	328
388	328
597	328
448	328

RESULT 9
US-10-09

; Sequence 53, Application US/10098841
 ; Publication No. US20020197679A1
 ; GENERAL INFORMATION:
 ;

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Asundi, Yandou
 APPLICANT: Xu, Chongjun
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yungqing
 APPLICANT: Wang, Yilan-Rui
 APPLICANT: Zhao, Qing A.
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhimei
 APPLICANT: Wehrman, Tom
 APPLICANT: Zhang, Jie
 APPLICANT: Qian, Xiaohong B.
 APPLICANT: Dimauc, Radcoje T.
 TITLE OF INVENTION: No. US20020197679a1 Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2
 CURRENT APPLICATION NUMBER: US/10/0/098,841
 CURRENT FILING DATE: 2002-03-13
 PRIOR APPLICATION NUMBER: 09/598,042
 PRIOR FILING DATE: 2000-06-20
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/468,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 331
 SOFTWARE: pc_pl_genes Version 1.0
 SEQ ID NO 53
 LENGTH: 1402
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (230)..(1075)
 US-10-098-841-53

Query Match : 23.2%; Score 377; DB 13; Length 1402;

Best Local Similarity 69.0%; Pred.No. 4.4e-103;
Matches 546; Conservative 0; Mismatches 240; Indels 5; Gaps 2;

QY	22	GGCGTTGGCGGGCGCGCGCGGCGCCCGGGAGAGGGCGGGCGCGGGAGCCCGATGCGCGGCA	81
	211	GGGAGAGGAGTCCGACCGCGGGCGCGGGAGCGTGTGCTTCGTTGCAACGGCAGCGG	270
Db	82	GGGAGAGCCGGAATGGCTTCCGCGGGAGCGCAAGCTCCAGAGCCCCCGCGCGAGGCC	141
QY	271	GAGGAGGGCGGCGCCGAAACCATGGCGGGGAAAGCGAGCAGCCCTTGAGAAACACACGCA	330
Db	142	GACCGTCCGCAACAGCGGCGCTTCTCGATAGGGGTGAGCGCGGCACTGCACGCGGGAAG	201
QY	331	GCCCAACGGCGCG--GAGCCCTTCTTATAGCGGTGACGGGGGAAACAGCTACGGCGAAG	388
Db	202	TGCAACCGTGTGAGAGAAATCATGGAATTCTGGGACAGAAACGAGGTGAAACACGCGCAG	261
QY	389	TCTTCCGTGTGCTTAAGATCGTGACGCTCTCGGGGACGAATGAGGTGAGCTATCGCCAG	448

Oy	26	CGAAGTGTGATCTCTGAGCCAGACAGGTTCTACAAAGTCTCGACGGACAGACAAG	321
Db	449	AGCAGGTGTCTATCTTGAGCCAGGATAGCTTTACCGTGTCCTTAACTCGAGACAGAG	508
Oy	322	GCCAAAGCCTTGAAAGACAGTACATTTTGCATCCATCCAGATGCTTGATATGATTTG	381
Db	509	GCCAAAGCCTGAGAGGCCAGTTCAACTTTGACACCCGGATGCTTTGACATGAACTC	568
Oy	382	ATGCACAGACCTCTGAGAACATCTGTGAGGGCAAAACGGTGGAGGATCCGACTATGAT	441
Db	569	ATTCTCAAAACACTCAAGAAATCACTGAAAGGAAACAGTCCAGATCCCGTGTATGAC	628
Oy	442	TTTGTACACACTCAAGGTTTACAGAGACAAGTGTCTACCTCGGAGAGTGGTTCTG	501
Db	629	TTTGTCTCCCATTTCCCGAAGGAGAGACAGTACTGTCTATCCCGAAGAGTGTGTCTC	688
Oy	502	TTTGAAGGAGATCCTTGCTGTCTTCTACAGCCAGAGATCCGGACATGTTTCACTTGCGCTC	561
Db	689	TTTGAAGGAGATCCTTGCGCTTCTTACTCCAGAGAGTACAGACCTGTTCCAGATTAACCTT	748
Oy	562	TTTGTGGACACCGAATCTCCGACGTCAGGCTGTCTCGAAGAGTTTCTCCGGACGTGCGC--	618
Db	749	TTTGTGATACAGATCCGACACACCGGCTCTCAGCGAGATTTAAAGGACATACAGGAG	808
Oy	619	CGAGGAGAGGACCTGGAGAGATTCTGACGACAGTACCACTCTGTGAAGCCGGCCCTTC	678
Db	809	AGAGGACAGGAGATCTTGAGCAGATTTTATCTCAGTACATTAAGTGTGTCAAGCTGTGCTTT	868
Oy	679	GAGAGATTTGCTGCTGCGACCAAGAAAGATGCCGATGTGATCATCCACGAGAGGTGGAC	738
Db	869	GAGGATTTGCTGTGCAACAAAGAAAGTATGTGATGATCATCCTTAAGAGTGCAGAT	928
Oy	739	AATATGTTGCCATCAACTGATCTGTGACGACATCCAGGACATTTGGAATGTGACATC	799
Db	929	AATCTGTGGCATCAACCTCATCTGTGACGACATCCAGGACATCTGGAATGAGGGCCC	988
Oy	799	TGCAATGGCA	809
Db	989	TCCAAACGGCA	999

RESULT 10
US-09-925-300-220

; Sequence 220, Application US/09925300
; Patent No. US20020151681A1

```

? GENERAL INFORMATION:
? APPLICANT: Craig Rosen,
? APPLICANT: Steve Ruben.
? TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
? FILE REFERENCE: PA101
? CURRENT APPLICATION NUMBER: US/09/925,300
? CURRENT FILING DATE: 2001-08-10
? PRIOR APPLICATION NUMBER: PCT/US00/05988
? PRIOR FILING DATE: 2000-03-08
? PRIOR APPLICATION NUMBER: 60/124,270
? PRIOR FILING DATE: 1999-03-12
? NUMBER OF SEQ ID NOS: 1890
? SOFTWARE: Patentn Ver. 2.0
? SEQ ID NO 220
? LENGTH: 1310
? TYPE: DNA
? ORGANISM: Homo sapiens
? JS-09-925-300-220

```

Query Match	Score	DB	Length
22.6%	367.2	10	1310

Best Local Similarity 68.6%; Pred.No. 3,9e-100;
Matches 550; Conservative 1; Mismatches 244; Indels 7; Gaps 3;

QY 14 CCACTCTGCGCTGGGCGGGCGCGCCGCGGAAAGGGCGGGGACCCGAT 73
140 CGGCTCGGAAAGCGGAGGAGTCCGACGCGGGGCGGGGAGCGTCTCCGTTTCG 199


```

QY 74 GCGGCGAGCGAGGCGGAGTGGCTTCGGGCGGAGCGAAGTCTCGAGAGCCCGCGC 133
    |||
Db 200 CACAGGCGAGCGGAGGAGGGGGCGCGCAACCATGGCGGGGACAGCGAGACCCCTGC 259
    |||
QY 134 CGAGAGCGGAGCGGCTCGGACCAAGCG--GCCCTTCGTATAGGGGTGAGCGCGGACTG 190
    |||
Db 260 AGAACCACACACAGCCCAACCGCGCGGAGCCCTTCCTATAGGCGCTCA-CGGGGGACAG 318
    |||
QY 191 CCAGCGGAGAGTGCACCGTGTGTGAGAGATCATGGATGTTGGGACAGAGACGAGGTGG 250
    |||
Db 319 CTAGCGGAGAGTCTTCCTGTGTGTGAGATGTCACACTCTGGGCGAGATGAGGTGG 378
    |||
QY 251 AACAGCGGACGCGAAGGTGTCTCTGAGCGAGACAGAGTTCACAGGTCTCTGACGG 310
    |||
Db 379 ACTATCGCAGAGAGGTGTGTCTCTGAGCGAGATGAGTTCACAGGTCTCTGACCT 438
    |||
QY 311 CAGAGCAGAGGCGGAGGCTTTGAAAGACAGTCAATTTTGACCATCCAGATGCTTTG 370
    |||
Db 439 CGGAGCAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCG 498
    |||
QY 371 ATTAATGATTTGATGACAGAGCTCTGAGAAATCTGTGAGGAGGAGGAGGAGGAGG 430
    |||
Db 499 ACAATGASTATTTCTCAAAACCTCAAGAAATCACTGAGGAGGAGGAGGAGGAGGAG 558
    |||
QY 431 CGACCTATGATTTGTGACACACTCAAGTTACAGAGACCAAGGTGTCTACCTCGCG 490
    |||
Db 559 CCGGTATGACTTTGTCTCCCACTTCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 618
    |||
QY 491 ACGGTGTTCTTTGAGGAGCATCTTGTGTCTTACAGCGAGAGATCCGGGACATGTTCC 550
    |||
Db 619 ACGGTGTTCTTTGAGGAGCATCTTGTGTCTTACAGCGAGAGATCCGGGACATGTTCC 678
    |||
QY 551 ACTGCGGCTCTGTGTGACACCGACTCGAGCGTGGGCTGTCTGAGAGATGTTCCGGG 610
    |||
Db 679 AGATGAGACTTTTGTGTGATACAGATGCGGACACCGGCTCTCACGAGATATTAAGG 738
    |||
QY 611 ACGTGGCG--GAGAGGAGGAGCTGAGCAGATTTGACGACGATCAACCACTTCTGTA 667
    |||
Db 739 ACATCAGGAGAGAGGAGGAGGATCTTGAGCAGATTTTATCTCAGTACATTAAGTTCG 798
    |||
QY 668 AGCCGCGCTTGAAGAGTTCGCTGCGGACAAAGAGATAGCGATGTATCATCCGAC 727
    |||
Db 799 AGCTGCTTTGAGAGAAATTCGCTTGCCCAACAAAGATAGCTGATGTATCATCCCTA 858
    |||
QY 728 GAGGAGTGCACATATGTTGCTCAATCAACTGATCGTACAGCATCCAGATATTTCTGA 787
    |||
Db 859 GAGGTGCAGATATCTGTGCTGCTCAACTCATCTGTCAGACATCCAGACATCTCTGA 918
    |||
QY 788 ATGTTGACATCTGCAATGGA 809
    |||
Db 919 ATGAGGCGGCTCCAAACGGCA 940
    |||

```

23

```

; NAME/KEY: misc.feature
; LOCATION: (1) -- (472)
; OTHER INFORMATION: n = A, T, C or G
US-09-918-995-24042

Query Match      13.6%; Score 221; DB 11; Length 472;
Best Local Similarity 73.0%; Pred. No. 4e-56;
Matches 284; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 159 GCCCTTCTGATAGGGGTGAGCGGCGGACATGCGAGCGGGAATCCAGCGTGTGTGAGA 218
    |||
Db 84 GCCCTTCTGATAGGGGTGAGCGGCGGAGACAGTACGCGGAAGTCTTCCGTGTGTCTAA 143
    |||
QY 219 GATCATGAGTTCCTGGGACAGAGAGAGTGAACAGCGGACGAGGAGGTGTCTCT 278
    |||
Db 144 GATGTGACACTCTCTGGGCGAGATGAGTGTGACTATCGCAGAGAGGATGATCTCT 203
    |||
QY 279 GAGCGAGACAGGTTCTTCAAGGTCTCTGACGCGAGAGCAGAGGCGGAGGCGGAGG 338
    |||
Db 204 GAGCGAGATAGCTTCTACGCTGCTTACCTCGAGAGAGGCGGAGGCGGAGGCGGAGG 263
    |||
QY 339 ACAGTCAATTTGACCATCAGATGCTTTGATTAATGATTTGATGACAGACTCTGAA 398
    |||
Db 264 CAGTTCACTTTGACCAAGCGGATGCTTTGACAACTGATCTCTCAAAACACTCAA 323
    |||
QY 399 GAACATGTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
    |||
Db 324 AGAATCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 383
    |||
QY 459 GTTACGAGAGACAGGAGTGTCTACCTCGGAGCGTGTCTTGTGAGGAGCATCTTGT 518
    |||
Db 384 GAAGAGAGAGACAGTACTGTATCCCGGAGAGTGTGTCTTTGAGGAGATCTCTGGC 443
    |||
QY 519 GTTCTACGCGAGAGATCCGGGACATGT 547
    |||
Db 444 CTCTACTCCAGAGAGTACGAGACTGT 472
    |||

RESULT 12
US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US2003087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pye, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1843
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1843

Query Match      13.2%; Score 214.8; DB 14; Length 447;
Best Local Similarity 73.0%; Pred. No. 2.9e-54;
Matches 276; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 159 GCCCTTCTGATAGGGGTGAGCGGCGGACATGCGAGCGGAGGAGTTCAGCGTGTGTGAGA 218
    |||

```



```

Db      70 GCCCTTCCTTAATAGCGCTCAGCGGGGAAACAGCTAGCGGCAAGTCTTCCGTGTGCTAA 129
Qy      219 GATCATGAGATGCTGGGAGAGACGAGTGGAAACAGCGGAGGGAAGGTGTCATCTT 278
Db      130 GATGTGACACTCTGGGGGAGAAATGAGTGAAGCTATTCGCAAGAGAGGTGTGATCTT 189
Qy      279 GAGCCAGAGAGAGTCTTCAAGAGTCTGACGGCAGACAGAGAGGCCAAGGCTTGAAGG 338
Db      190 GAGCCAGAGATAGCTTCAACCTGTCTTACCTCGAGAGAGAGGCCAAGGCTTGAAGG 249
Qy      339 ACAGTACAATTTTGGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTGAA 398
Db      250 CCAATTCACTTGTGACCAACCGGATGCTTTGACATGAATCAATCTCAAAAACATCAA 309
Qy      399 GAACATGTGAGGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGTGACACTGAA 458
Db      310 AGAATACTGAAAGGAAACAGTCCAGATCCCGTATGACTTGTCTCCCAATCCG 369
Qy      459 GTTACAGAGACCAAGGTGCTACCTGCGAGAGCTGTTCTGTTGAGGGCATCTTGT 518
Db      370 GAAGGAGAGACAGTACTGTCTATCCGAGAGAGTGTGCTTTGAAGAACTCTGCC 429
Qy      519 GTTCTACAGCCAGAGAT 536
Db      430 CTCTACTCCAGAGAGT 447

```

RESULT 13

```

US-09-918-995-23923
; Sequence 23923, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23923
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

```

```

Query Match      13.0%; Score 210.4; DB 11; Length 455;
Best Local Similarity 72.8%; Pred. No. 6.3e-53;
Matches 271; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```

```

Qy      159 GCCCTTCCTTAATAGCGGTGAGCGGCACTGCCAGGGAAGTGCACCGTGTGAGAA 218
Db      84 GCCCTTCCTTAATAGCGGTGAGCGGGAACAGCTAGGCGCAAGCTTCCGTGTGCTAA 143
Qy      219 GATCATGAGATGCTGGGAGAGACGAGTGGAAACAGCGGAGCGAAGGTGTCATCTT 278
Db      144 GATGTGACACTCTGGGGGAGAAATGAGTGAAGCTATTCGCAAGAGAGGTGTGATCTT 203
Qy      279 GAGCCAGAGAGAGTCTTCAAGAGTCTGACGGCAGACAGAGGCCAAGGCTTGAAGG 338
Db      204 GAGCCAGAGATAGCTTACCGGTCTTACCTGGAGACAGAGGCCAAGGCCCTGAAGG 263
Qy      339 ACAGTACAATTTTGGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTGAA 398
Db      264 CCAATTCACTTGTGACCAACCGGATGCTTTGACATGAATCAATCTCAAAAACATCAA 323
Qy      399 GAACATGTGAGGGGCAAAACGTTGAGAGTCCGACCTATGATTTTGTGACACTGAA 458

```

```

Db      324 AGAATCAGTGAAGGGAAGAAACAGTCCAGATCCCGGTGATGACTTTGTCCCATTCGCG 383
Qy      459 GTTACAGAGACCAAGGTGCTTACCTGCGAGCGTGTCTTGTGAGGGCATCTTGT 518
Db      384 GAAGGAGAGACAGTACTGTCTATCCCGAGAGCGTGTGCTTGTGAAGGATCTCGGC 443
Qy      519 GTTCTACAGGCA 530
Db      444 CTCTACTCCCA 455

```

RESULT 14

```

US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A192171.1, EVALUATE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUATE 4.00e-26
; OTHER INFORMATION: NT HIT: g114783235, EVALUATE 1.00e-100
US-10-029-386-15883

```

```

Query Match      11.4%; Score 185; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.9e-45;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      745 GTTGCATCAACTGATCGTGACAGACATCCAGAGCAATTTGTGATGTGACATCTGCAAA 804
Db      3 GTTGCATCAACTGATCGTGACAGACATCCAGAGCAATTTGTGATGTGACATCTGCAAA 62
Qy      805 TGGCAGCGAGAGAGGTCCAAATGGCGGAGCTCAACGCAACCTTTCTAGCCAGGGGAC 864
Db      63 TGGCAGCGAGAGAGGTCCAAATGGCGGAGCTCAACGCAACCTTTCTAGCGGGGAC 122
Qy      865 CACCTGGAGATGCTACCTCTGGCAAAAGGTCAATTTGAGTCCAGAGAGAGAGAGAG 924
Db      123 CACCTGGAGATGCTACCTCTGGCAAAAGGTCAATTTGAGTCCAGAGAGAGAGAGAG 182
Qy      925 TGAGG 929
Db      183 TGAGG 187

```

```

RESULT 15
US-10-251-186-14
; Sequence 14, Application US/10251186
; Publication No. US20030180745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom

```

```
; APPLICANT: Zhang, Jie
; APPLICANT: Dimanac, Radoje T.
; TITLE OF INVENTION: No. US20030180745A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 789CIP2CDV1
; CURRENT APPLICATION NUMBER: US/10/251,186
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PL_FL_genes Version 2.0
; SEQ ID NO 14
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(1909)
US-10-251-186-14
```

```
Query Match      10.8%; Score 175.6; DB 12; Length 2058;
Best Local Similarity 58.4%; Pred.No. 4.9e-42;
Matches 326; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

QY      236 GACGAAACGAGGTGGAAACAGCGGCGAAGCGGAAAGTGTCATCTTGAGCCAGACAGGTTCT 295
Db      609 GAATGATATGAGAGGCCCTGATGTGCTCGGTGGTGTCTTGTCTCCATGATGATCTCTCT 668

QY      296 ACAAGGTCTTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 355
Db      669 ACAAGGTCTTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728

QY      356 ATCCAGATGCTTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 415
Db      729 ACCCAGATGCTTGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 788

QY      416 AAACGGTGAAGTGGCCGACCTATGATTTGTGACACACTCAAGTTTACAGAGACCAAGG 475
Db      789 AGAGTGTCAAGTGGCCGACCTATGATTTGTGACACACTCAAGTTTACAGAGACCAAGG 848

QY      476 TGTCTACCCCTGCGGAGCGTGGTTCGTTTGAAGGATCTTGATCTACAGCCAGGAGA 535
Db      849 CACTGTATGTGCAAAACGTCATCTTTGAGGGATCATGCGCTTGTCTGACAGAGAC 908

QY      536 TCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACCGACTCCGAGCTGAGCTGTCTC 595
Db      909 TGTGTGAGCTCTGTGACATGAAGATCTTTGTGACACAGACTCCGACATCCGCTGGTAC 968

QY      596 GAAAGTTCTCCGGAGCGTGG--CCGAGGAGAGAGACTTGAGCAGATTCTGACGCACT 652
Db      969 GCGGCTGCGCGCGGAGACATGATGAGCGCGGCGGAGCATGAGAGGTTCATCAAGCAGT 1028

QY      653 ACACCACTTGTGAGAGCGGCGCTTTCGAGAGTTCGCTGCGCAAAAGATATGCGG 712
Db      1029 ACAACAAGTTGTCAAGCTCTCTTCGACGATACATCAAGCCCAACATGCGCTGGCAG 1088

QY      713 ATGTGATCATCCAGAGAGATGAGCAATATGTTGCCATCAACCTGATCGTGCAGACA 772
Db      1089 ACATCGTGTCCCAAGAGGAGCGGCAACAGGTGGCCATGACCTGATTGTGACAGCAG 1148

QY      773 TCCAGGACATTCTGAATG 790
Db      1149 TGCAAGCCAGCTGAGG 1166
```

Search completed: November 25, 2003, 02:18:49
Job time : 598.791 secs

XX 10-JAN-2002.
PD 28-JUN-2001; 2001WO-US21063.
XX 30-JUN-2000; 2000US-216503P.
XX (MILL-) MILENNIUM PHARM INC.
XX Glucksmann MA;
XX WPI; 2002-140091/18.
DR P-PSDB; AAE16592.
XX
XX New isolated human uridine kinase family polypeptide 57658, useful for
PT treating hematopoietic neoplastic disorders and disorders of neurons,
PT heart and blood vessels -
XX
PS Claim 1; Fig 1a; 103pp; English.
XX
XX The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as hematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.
XX
SQ Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other:
Query Match 100.0%; Score 1624; DB 24; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGGGTCTGCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGCGGC 60
DB 1 GTGGGGTCTGCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGGCGCGGC 60
QY GCGGGAGCCGATCGCGGGAGCGGAGCGCGATGCTTGCGGGAGCGGAGCTGC 120
DB 61 GCGGGAGCCGATCGCGGGAGCGGAGCGCGATGCTTGCGGGAGCGGAGCTGC 120
QY 121 GAGAGCCCGCGCGCGGAGCGCGATCGCGACCGAGCGCGCTTCTGATGAGGATGAGC 180
DB 121 GAGAGCCCGCGCGCGGAGCGCGATCGCGACCGAGCGCGCTTCTGATGAGGATGAGC 180
QY 181 GCGCGCACTGCGACCGGGAATCGACCGTGTGTGAGAAAGATCATGAGATTGCTGGGACAG 240
DB 181 GCGCGCACTGCGACCGGGAATCGACCGTGTGTGAGAAAGATCATGAGATTGCTGGGACAG 240
QY 241 AAGAGGTGAAACAGCGGCGAGCGGAAGGTGCTCATCTCGAGCCGAGACAGGTTTACAAG 300
DB 241 AAGAGGTGAAACAGCGGCGAGCGGAAGGTGCTCATCTCGAGCCGAGACAGGTTTACAAG 300
QY 301 GTCTCGACGCGAGACGAGAAAGCGCTTGAAGAGACAGTAAATTTGACATCA 360
DB 301 GTCTCGACGCGAGACGAGAAAGCGCTTGAAGAGACAGTAAATTTGACATCA 360
QY 361 GATCCCTTTGATTAATGATTTGATGACAGACTCTGAAGAACATCTGAGAGGCGAAACG 420
DB 361 GATCCCTTTGATTAATGATTTGATGACAGACTCTGAAGAACATCTGAGAGGCGAAACG 420
QY 421 GTGAGAGGTGCGGACCTATGATTTGTGACACATCAAGGTTACAGAGACCAAGGTGTC 480

DB 421 GTGAGAGGTGCGGACCTATGATTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTC 480
QY 481 TACCTCGGAGCGTGGTCTGTTTGAAGGCACTTGGTGTCTACAGCAGAGATCCG 540
DB 481 TACCTCGGAGCGTGGTCTGTTTGAAGGCACTTGGTGTCTACAGCAGAGATCCG 540
QY 541 GACATGTTCACTGCGGCTTCTGTTGACACCGAATCCGAGCTCAGGCTCTCGAAGA 600
DB 541 GACATGTTCACTGCGGCTTCTGTTGACACCGAATCCGAGCTCAGGCTCTCGAAGA 600
QY 601 GTTCTCGGAGCGTGGCGGAGGAGGAGCCTGAGACAGATTCTGACGAGTACACACC 660
DB 601 GTTCTCGGAGCGTGGCGGAGGAGGAGCCTGAGACAGATTCTGACGAGTACACACC 660
QY 661 TTGTAAGCGGCGCTTGAAGAGTTCCTGCTGCGGACAAAGATATGCGCATGATC 720
DB 661 TTGTAAGCGGCGCTTGAAGAGTTCCTGCTGCGGACAAAGATATGCGCATGATC 720
QY 721 ATCCACAGAGATGAGCAATATGTTGCAATCACTGATGCGAGACATCCAGAC 780
DB 721 ATCCACAGAGATGAGCAATATGTTGCAATCACTGATGCGAGACATCCAGAC 780
QY 781 ATTCTGAATGATGACATCTGCAAAATGAGCAGGAGGCTCCATGAGGAGCTCAAG 840
DB 781 ATTCTGAATGATGACATCTGCAAAATGAGCAGGAGGCTCCATGAGGAGCTCAAG 840
QY 841 CGGACCTTTTCTGAGCGAGGAGCAACCTGGAGTCTGACCTTGGCAACCGTCAAT 900
DB 841 CGGACCTTTTCTGAGCGAGGAGCAACCTGGAGTCTGACCTTGGCAACCGTCAAT 900
QY 901 TTGGAATCCAGACGACAGACCCCTGAGGCGTCCGAGCTCAGGCGAGTCTCCGCG 960
DB 901 TTGGAATCCAGACGACAGACCCCTGAGGCGTCCGAGCTCAGGCGAGTCTCCGCG 960
QY 961 CGGCAATGTTGTTCAAGGACCTGAGCTGGGAGCGCCACCACTGCTTCTCTC 1020
DB 961 CGGCAATGTTGTTCAAGGACCTGAGCTGGGAGCGCCACCACTGCTTCTCTC 1020
QY 1021 GCGCAACCCAGGAGGAGTGTAGACGAGGCTTCTCTCACTGAGAGTGAACCTCAGA 1080
DB 1021 GCGCAACCCAGGAGGAGTGTAGACGAGGCTTCTCTCACTGAGAGTGAACCTCAGA 1080
QY 1081 TGTGTACTCAGACTCAACTGTGTGAGACATGACAGGGTCTGAGGTTTCAAGCAC 1140
DB 1081 TGTGTACTCAGACTCAACTGTGTGAGACATGACAGGGTCTGAGGTTTCAAGCAC 1140
QY 1141 TTAGGCTCGTTGCGGTTTAAAGATCCCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
DB 1141 TTAGGCTCGTTGCGGTTTAAAGATCCCTAGGTCACTGAGAAATGCCAGAAATGTGCA 1200
QY 1201 GGAAGCTTGGAGAGCTTCTGTGAGAAATGAGGACATTTTGGGAAATTTGAGAGAC 1260
DB 1201 GGAAGCTTGGAGAGCTTCTGTGAGAAATGAGGACATTTTGGGAAATTTGAGAGAC 1260
QY 1261 AGGCTAGACACTGCGCGGCTGATGTTTGTGACAGTAAACCAAGTGGAGAGAGTT 1320
DB 1261 AGGCTAGACACTGCGCGGCTGATGTTTGTGACAGTAAACCAAGTGGAGAGAGTT 1320
QY 1321 TTTTCAAGTCTGATCTGTTCTTACACACTCACTACACATTACTCAAAATTTTGTGACA 1380
DB 1321 TTTTCAAGTCTGATCTGTTCTTACACACTCACTACACATTACTCAAAATTTTGTGACA 1380
QY 1381 AGTACTTCTCTTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAATACA 1440
DB 1381 AGTACTTCTCTTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAATACA 1440
QY 1441 CAAGGCTGTTGTTGCTTCAAACTTAATTTATGACCAAGGCTTTGAGCTCAGCTG 1500
DB 1441 CAAGGCTGTTGTTGCTTCAAACTTAATTTATGACCAAGGCTTTGAGCTCAGCTG 1500
QY 1501 GCTTACACGATATGAGGAGCCACTGAGGAGATTTTCCCTTCTGCTTGAAG 1560

Db 967 TGTGTTCAGGAGCATGAGCGCTGGGGAGGCCCAACCAACCACTGCTCTCTGGCGC 1026
Qy 1026 ACCCCAGGGAGGTGTAGACGCGAGCGCTTCTCACTCAGAGTGGAACTCAGATGTGT 1085
Db 1027 ACCCCAGGGAGGTGTAGACGCGAGCGCTTCTCTCACTCAGAGTGGAACTCAGATGTGT 1086
Qy 1086 CAGTCAGACTCAACTGTGTGGGACATGACAGCGGCTTCGAGGTTTTCAGCCACTTAAAG 1145
Db 1087 CACTCAACTCAACTGTGTGGGACATGACAGCGGCTTCGAGGTTTTCAGCCACTTAAAG 1146
Qy 1146 CTGCTGTCGGTTTAAAGATCCCTCTAGATCACTAGAAATGCCAGAAATGTGCAGAG 1205
Db 1147 CTGCTGTCGGTTTAAAGATCCCTCTAGATCACTAGAAATGCCAGAAATGTGCAGAG 1206
Qy 1206 CTTGGGAGGCTTCTGTGAGGAATGTGAGGCATTTATTTGGGAAATTTGAGAGACACCT 1265
Db 1207 CTTGGGAGGCTTCTGTGAGGAATGTGAGGCATTTATTTGGGAAATTTGAGAGACACCT 1266
Qy 1266 AGACACTGGGCTGGCTGATGTTTGTGACAGTGAACCAACAGTGGGAGAGATTTTTC 1325
Db 1267 AGACACTGGGCTGGCTGATGTTTGTGACAGTGAACCAACAGTGGGAGAGATTTTTC 1326
Qy 1326 CAGTCTGATCTGTTCTTACACACTCACACATTAACCTAAAGTTTGTGAACAAGTAC 1385
Db 1327 CAGTCTGATCTGTTCTTACACACTCACACATTAACCTAAAGTTTGTGAACAAGTAC 1382
Qy 1386 TTTCTTTTTCATGTTTACATGTCCTCATGTTTCTGTTTCTGTTTTCATACACAGG 1445
Db 1383 TTTCTTTTTCATGTTTACATGTCCTCATGTTTCTGTTTCTGTTTTCATACACAGG 1442
Qy 1446 CTGCTGTCGGCTTCAAAACCTTAATTTATGACCCAGTGGTGTGAGTCCAGCGTGGCTA 1505
Db 1443 CTGCTGTCGGCTTCAAAACCTTAATTTATGACCCAGTGGTGTGAGTCCAGCGTGGCTA 1502
Qy 1506 CACGGATATGAGGAGCACTGAGGAGATGTTTCCCTTGTGCTTAAAGGAGAG 1565
Db 1503 CACGGATATGAGGAGCACTGAGGAGATGTTTCCCTTGTGCTTAAAGGAGAG 1562
Qy 1566 GAAGCGAGGCGGAGTCCCTTGAAGACACCCAGCATCACACCGGCTTGTGGGGCCAG 1624
Db 1563 GAAGCGAGGCGGAGTCCCTTGAAGACACCCAGCATCACACCGGCTTGTGGGGCCAG 1621

RESULT 3
AA160444
ID AA160444 standard; cDNA; 2152 BP.
XX
AC AA160444;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4433.
XX
KW Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QH, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41288.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4433; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM36642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
SQ Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Query Match 97.6%; Score 1585.8; DB 22; Length 2152;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 36 CGCGGCGCGCGGGAAGGGGCGGCGCGGCGGACCCGATCGCGGAGCGGAGCGCGAGAT 95
Db 1 CGCGGCGCGCGGGAAGGGGCGGCGCGGCGGACCCGATCGCGGAGCGGAGCGCGAGAT 60
Qy 96 GCGTTGGCGGGAGGCGGAAGACTGCGAGAGCCCGCGCGCGGAGCCGATCCGACCA 155
Db 61 GCGTTGGCGGGAGGCGGAAGACTGCGAGAGCCCGCGCGCGGAGCCGATCCGACCA 120
Qy 156 GCGGCGCTTCGTATGAGGGGTGAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGA 215
Db 121 GCGGCGCTTCGTATGAGGGGTGAGCGCGGCACTGCGCAGCGGGAAGTCGACCGTGTGA 180
Qy 216 GAAGATCATGAGATTCTGTGGACAGAACGAGGTGGAACGCGGACGCGAAGGTGTCTAT 275
Db 181 GAAGATCATGAGATTCTGTGGACAGAACGAGGTGGAACGCGGACGCGAAGGTGTCTAT 240
Qy 276 CTTGAGCCAGGACAGGTTCTTCAAGGTCTGTACGCGGAGGCGGAAGGCCAGGCTTGA 335
Db 241 CTTGAGCCAGGACAGGTTCTTCAAGGTCTGTACGCGGAGGCGGAAGGCCAGGCTTGA 300
Qy 336 AGGACGTACAATTTTGACATCCAGATGCTTTGTAATGATTTGATGACAGACTCT 395
Db 301 AGGACGTACAATTTTGACATCCAGATGCTTTGTAATGATTTGATGACAGACTCT 360
Qy 396 GAAGAATCATGTGAGAGGCAAAAAGGTGAGGTGCGGACTATGATTTGTGACACTC 455
Db 361 GAAGAATCATGTGAGAGGCAAAAAGGTGAGGTGCGGACTATGATTTGTGACACTC 420
Qy 456 AAGTTACAGAGACACAGGTGTCTTACCTTGGGACGTTGTTCTTTGAGGGCATCTT 515
Db 421 AAGTTACAGAGACACAGGTGTCTTACCTTGGGACGTTGTTCTTTGAGGGCATCTT 480

OY	114	AACATCGCAGAGCGCCCGCGCCGAGAGCCGACCGTCCGACACAGCGGCGCTTCTGTAAG	173
Db	61	AACATCGCAGAGCGCCCGCGCGGAGCCGACCGTCCGACACAGCGGCGCTTCTGTAAG	120
OY	174	GGTGAAGCGCGGCACTGCGCAGCGGGAAGTGCACCTGTGTGAAGAAATCATGAGATTGGCT	233
Db	121	GGTGAAGCGCGGCACTGTCAGCGGGAAGTGCACCGTGTGTGAAGAAATCATGAGATTGGCT	180
OY	234	GGGACAGAAAGAGTGGAAACAGCGCAGCGGGAAGTGTCAATCTGAGCCAGGACAGATT	293
Db	181	GGGACAGAAAGAGTGGAAACAGCGGCGAGCGGGAAGTGTCAATCTGAGCCAGGACAGATT	240
OY	294	CTAACAAAGTCTGACCGGCAAGACAGAAAGGCCAAGGCTTGAAAGGACAGTCAATTTTGA	353
Db	241	CTAACAAAGTCTGACCGGCAAGACAGAAAGGCCAAGGCTTGAAAGGACAGTCAATTTTGA	300
OY	354	CCATCCAGATGCGCTTGTATATGATTTGATGACAGAGACTGTGAAGAAATCGTGGAGG	413
Db	301	CCATCCAGATGCGCTTGTATATGATTTGATGACAGAGACTGTGAAGAAATCGTGGAGG	360
OY	414	CAAAACGTTGAGAGTGCAGCCTATAGATTGTGACACTCAAGGTTACAGAGACAC	473
Db	361	CAAAACGTTGAGAGTGCAGCCTATAGATTGTGACACTCAAGGTTACAGAGACAC	420
OY	474	GGTGTCTTACCTTCGCGAGAGTGTCTGTGTTGAGGCACTTGTGTCTTACACCGAGA	533
Db	421	GGTGTCTTACCTTCGCGAGAGTGTGTCTGTTGAGGCACTTGTGTCTTACACCGAGA	480
OY	534	GATCCGGGACATGTTTCACTGCGGCTTTCGTGACACCGAGCTCGAGGTCAGGCTGTC	593
Db	481	GATCCGGGACATGTTTCACTTGCCTCTTTCGTGACACCGAGCTCGAGGTCAGGCTGTC	540
OY	594	TGGAAGAGTTCCTCCGGAGCGTGCGCCGAGGAGAGGACCTGAGACAGATTCTGACGCA	653
Db	541	TGGAAGAGTTCCTCCGGAGCGTGCGCCGAGGAGAGGACCTGAGACAGATTCTGACGCA	599
OY	654	CACCAACCTTCGTGAAGACCGGCTTTCGAGGAGTTCCTCTCCGCACAAAGAATGCGCA	713
Db	600	CACCAACCTTCGTGAAGACCGGCTTTCGAGGAGTTCCTCTCCGCACAAAGAATGCGCA	659
OY	714	TGTGATCATCCACGAGGAGTGGACAAATATGTTGCATCAACTGATGTGTGAGACAT	773
Db	660	TGTGATCATCCACGAGGAGTGGACAAATATGTTGCATCAACTGATGTGTGAGACAT	719
OY	774	CCAGACATTTCTGAATGTGATCATCTGCAATGAGCACCCAGAGAGGTCCAATGAGCGAG	833
Db	720	CCAGACATTTCTGAATGTGATCATCTGCAATATGAGCACCGAGAGGTCCAATGAGCGAG	779
OY	834	CTACAAAGGAGACCTTTTCTGAGCCAGGAGGACCACTTGGAGTGTGACTCTTGCCAAACG	893
Db	780	-TACAAAGGAGACCTTTTCTTRACCAGGAGGACCACTTGGAGTGTGACTCTTGCCAAACG	838
OY	894	GTCACATTTGAGTCCAGACAGACACCCACGAGGAGCTGCAGAGCTCAGGAGAGTGC	953
Db	839	GTCACATTTGAGTCCAGACAGACACCCACTGAGGAGCTGCAGAGCTCAGGAGAGTGC	898
OY	954	TCCGCGCCCGCATGTGTGTTCAAGGACTGA-6CTTGGGACGCCACCAACCAACCACTGC	1017
Db	899	TCCGCGCCCGCATGTGTGTTCAAGGACTGA-6CTTGGGACGCCACCAACCAACCACTGC	958
OY	1013	TTCTCTCTGGGCGCACCCCAAGGAGTGTGAGCAGAGGCTTCTCACTCAGAGTGA	1077
Db	959	TTCTCTCTGGGCGCACCCCAAGGAGTGTGAGCAGAGGCTTCTCACTCAGAGTGA	1018
OY	1073	AACCTAGATGTCTCACTGACATCAACTTGTGGGACCTGACAGGCGTTCTTGAGGTT	1133
Db	1019	AACCTAGATGTCTCACTGACATCAACTTGTGGGACCTGACAGGCGTTCTTGAGGTT	1077
OY	1133	TCAGCACTTAGGCTGTGTCGGTTTAAAGATCCCTTAGTCACTGAGAAATGACAG	1197
Db	1079	TCAGCACTTAGGCTGTGTCGGTTTAAAGATCCCTTAGTCACTGAGAAATGACAG	1133
OY	1193	AATGTGCAGGAAGCTCGGAGGCTTCTGTGAGAAATGTAGGACATTAATGGGAAATT	1257

Db	1139	AATGTGACGAGAGCCTGGAGAGCTTCTGTGAGGAATGTGAGGACCATTAATTTGGGGAATT	1196
Oy	1253	GAGAGAGACAGCTTAGACACTGGCTGGCTGTGATGTTTTTTTGTAGACAGTGAACCAACAGTGGG	1312
Db	1199	GAGAGAGACAGCCTTAGACACTGGCTGGCTGTGATGTTTTTTTGTAGACAGTGAACCAACAGTGGG	1255
Oy	1313	AGAGAGTTTTTTTCAGTGTGATCTGGTTCCTTACACACTCACAACATTAACTCAAAAGTTT	1372
Db	1259	AGAGAGTTTTTTTCAGTGTGATCTGGTTCCTTACACACTCACAACATTAACTCAAAAGTTT	1314
Oy	1373	TGTGAACACAGTACTTTCCTTTTAA	1397
Db	1315	TGTGAACACAGTACTTTCCTTTTAA	1339
RESULT 5			
AA158658	AA158658 standard; cDNA; 1288 BP.		
XX	AA158658;		
XX	22-OCT-2001	(first entry)	
XX	Human polynucleotide SEQ ID NO 861.		
DE	Human polynucleotide SEQ ID NO 861.		
XX	Human; neurotrophic; immunosuppressant; cytoskeletal; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukemia; ss.		
XX			
OS	Homo sapiens.		
XX			
XX	WO200153312-A1.		
PN			
PD	26-JUL-2001.		
XX			
PF	26-DEC-2000; 2000WO-US34263.		
XX			
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou F, Goodrich R, Drmanac RT;		
XX			
XX	WPI; 2001-442253/47.		
DR	P-PSDB; AAM39502.		
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
PS	Claim 1; SEQ ID NO 861, 10078bp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA16136) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,		
CC	immunopressant and cytoskeletal activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		

PR	11.-SEP.-2000	2000US.-0229344
PR	01.-SEP.-2000	2000US.-0229345
PR	05.-SEP.-2000	2000US.-0229349
PR	05.-SEP.-2000	2000US.-0229351
PR	06.-SEP.-2000	2000US.-0230437
PR	06.-SEP.-2000	2000US.-0230438
PR	06.-SEP.-2000	2000US.-0231242
PR	08.-SEP.-2000	2000US.-0231243
PR	08.-SEP.-2000	2000US.-0231244
PR	08.-SEP.-2000	2000US.-0231413
PR	08.-SEP.-2000	2000US.-0231414
PR	08.-SEP.-2000	2000US.-0231480
PR	08.-SEP.-2000	2000US.-0233081
PR	12.-SEP.-2000	2000US.-0231968
PR	14.-SEP.-2000	2000US.-0232397
PR	14.-SEP.-2000	2000US.-0232398
PR	14.-SEP.-2000	2000US.-0233398
PR	14.-SEP.-2000	2000US.-0233399
PR	14.-SEP.-2000	2000US.-0233400
PR	14.-SEP.-2000	2000US.-0233401
PR	14.-SEP.-2000	2000US.-0233654
PR	14.-SEP.-2000	2000US.-0233653
PR	14.-SEP.-2000	2000US.-0233065
PR	21.-SEP.-2000	2000US.-0234223
PR	21.-SEP.-2000	2000US.-0234274
PR	25.-SEP.-2000	2000US.-0234997
PR	25.-SEP.-2000	2000US.-0234998
PR	26.-SEP.-2000	2000US.-0235484
PR	27.-SEP.-2000	2000US.-0235834
PR	27.-SEP.-2000	2000US.-0235836
PR	29.-SEP.-2000	2000US.-0236367
PR	29.-SEP.-2000	2000US.-0236368
PR	29.-SEP.-2000	2000US.-0236369
PR	29.-SEP.-2000	2000US.-0236370
PR	02.-OCT.-2000	2000US.-0237037
PR	02.-OCT.-2000	2000US.-0237038
PR	02.-OCT.-2000	2000US.-0237039
PR	13.-OCT.-2000	2000US.-0237040
PR	13.-OCT.-2000	2000US.-0241878
PR	20.-OCT.-2000	2000US.-0241877
PR	20.-OCT.-2000	2000US.-0241878
PR	20.-OCT.-2000	2000US.-0241809
PR	20.-OCT.-2000	2000US.-0241826
PR	01.-NOV.-2000	2000US.-0244617
PR	08.-NOV.-2000	2000US.-0246474
PR	08.-NOV.-2000	2000US.-0246475
PR	08.-NOV.-2000	2000US.-0246476
PR	08.-NOV.-2000	2000US.-0246477
PR	08.-NOV.-2000	2000US.-0246478
PR	08.-NOV.-2000	2000US.-0246523
PR	08.-NOV.-2000	2000US.-0246524
PR	08.-NOV.-2000	2000US.-0246525
PR	08.-NOV.-2000	2000US.-0246611
PR	17.-NOV.-2000	2000US.-0249207
PR	17.-NOV.-2000	2000US.-0249208
PR	17.-NOV.-2000	2000US.-0249209
PR	17.-NOV.-2000	2000US.-0249210
PR	17.-NOV.-2000	2000US.-0249211
PR	17.-NOV.-2000	2000US.-0249212
PR	17.-NOV.-2000	2000US.-0249213

Query Match	Best Local Similarity	Score	DB	Length	9732;
Matches	880;	Conservative	0;	Mismatches	0;
				Indels	Gaps
QY	745	GTTCGATTCACACCGATCGTCGACACATCCAGAGCATTCGATGGTGACATCTGCAGAA	804		
DB	3901	GTTCGATTCACACCGATCGTCGACACATCCAGAGCATTCGATGGTGACATCTGCAGAA	3960		
QY	805	TGGCAGCCGAGAGGGTCCATGGCGGAGCTACAAGCGAGACCTTTTCTGACCGAGGAGC	864		
DB	3961	TGGCAGCCGAGAGGGTCCATGGCGGAGCTACAAGCGAGACCTTTTCTGACCGAGGAGC	4020		

XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer -
XX
PS Disclosure; SEQ ID NO 2804; 321pp; English.
XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals
CC or humans. The nucleic acid: stimulates re-vascularisation of ischemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
CC used to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts; prevents skin aging due to sunburn by
CC stimulating keratinocyte growth; prevents hair loss, since FGF family
CC members activate hair-forming cells and promotes melanocyte growth;
CC stimulates growth and differentiation of hematopoietic cells and bone
CC marrow cells when used in combination with other cytokines; maintains
CC organs before transplantation or for supporting cell culture of primary
CC tissues; induces tissue of mesodermal origin to differentiate in early
CC embryos; increases or decreases the differentiation or proliferation of
CC embryonic stem cells, besides, haematopoietic lineage; modulates
CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's metabolic or physical state by influencing biorhythms,
CC circadian rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docid=20020147140.
CC
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Query Match 54.2%; Score 880; DB 25; Length 9732;
Best Local Similarity 100.0%; Pred. No. 1.9e-217;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCCATCAACCTGATCGTGACACATCCAGACATTCTGAATGTGACATTGGCAA 804
Db 3901 GTTGCCATCAACCTGATCGTGACACATCCAGACATTCTGAATGTGACATTGGCAA 3960

QY 805 TGGCACCAGAGAGGCTCAATGGCGGAGCTACAAGGGGACCTTTCTGAGCCAGGGGAC 864
Db 3961 TGGCACCAGAGAGGCTCAATGGCGGAGCTACAAGGGGACCTTTCTGAGCCAGGGGAC 4020

QY 865 CACCTTGGATGTGACCTTGCGMAACGGTCACTTTGAGTCCAGACAGACCCAC 924
Db 4021 CACCTTGGATGTGACCTTGCGMAACGGTCACTTTGAGTCCAGACAGACCCAC 4080

QY 925 TGAGGGGCTGCCAGGCTCAGGGGAGGTCTCCGCCCGGCACTGTGTTCAGGGACTGAG 984
Db 4081 TGAGGGGCTGCCAGGCTCAGGGGAGGTCTCCGCCCGGCACTGTGTTCAGGGACTGAG 4140

QY 985 CTTGGGAGCGGCCACCAACCCACTGCTTCTGTGGGCGACCCCGGGGAGGTAGC 1044
Db 4141 CTTGGGAGCGGCCACCAACCCACTGCTTCTGTGGGCGACCCCGGGGAGGTAGC 4200

QY 1045 AGCGAGGCTTCTCACTCAGAGTGAATCACTAGATGTCTCACTCAGACTCACTTGTCT 1104
Db 4201 AGCGAGGCTTCTCTCACTCAGAGTGAATCACTAGATGTCTCACTCAGACTCACTTGTCT 4260

QY 1105 GGGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTAGGCTCTGCGTTTAAAGAT 1164

Db 4261 GGGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTAGGCTCTGCGTTTAAAGAT 4320

QY 1165 CCTCTAGGTCATGAGAAATGCCACAGATGCGAGGAAGCCTGGAGGCTCTGTGAG 1224
Db 4321 CCTCTAGGTCATGAGAAATGCCACAGATGCGAGGAAGCCTGGAGGCTCTGTGAG 4380

QY 1225 GAATGTAGGACATTAATTTGGGAAATTTAGAGACAGCTTAGACACTGGCTGGCTGAT 1284
Db 4381 GAATGTAGGACATTAATTTGGGAAATTTAGAGACAGCTTAGACACTGGCTGGCTGAT 4440

QY 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 1344
Db 4441 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTA 4500

QY 1345 CACACTCACACATTAACCTCAAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTA 1404
Db 4501 CACACTCACACATTAACCTCAAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTA 4560

QY 1405 CATGCTCATGTTTCTGTTTCTGTTTCATTAACAAGGCTGTTGTGCTTACAAAC 1464
Db 4561 CATGCTCATGTTTCTGTTTCTGTTTCTGTTTCATTAACAAGGCTGTTGTGCTTACAAAC 4620

QY 1465 CTAAATTCATGACCCAGTGTGTCAGTCCAGCTGCTTACACGGATATGGGAGCCAC 1524
Db 4621 CTAAATTCATGACCCAGTGTGTCAGTCCAGCTGCTTACACGGATATGGGAGCCAC 4680

QY 1525 TGAGGGATGTTTCCCCCTGCTTGACCTTAAAGGACAGAGAGCGGATGCCCCCT 1584
Db 4681 TGAGGGATGTTTCCCCCTGCTTGACCTTAAAGGACAGAGAGCGGATGCCCCCT 4740

QY 1585 GGAAGCACCAGCATCACACCAGGCTTGTGCGGGCCAG 1624
Db 4741 GGAAGCACCAGCATCACACCAGGCTTGTGCGGGCCAG 4780

RESULT 8
AAL36440
ID AAL36440 standard; DNA; 19125 BP.
XX
AC AAL36440;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;
KW valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
PN MO20015367-A1.
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205615.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR	07-JUL-2000	2000US-0216680
PR	11-JUL-2000	2000US-0217487
PR	11-UTL-2000	2000US-0217496
PR	14-UTL-2000	2000US-0218290
PR	26-UTL-2000	2000US-0220963
PR	26-UTL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225568
PR	14-AUG-2000	2000US-0225568
PR	14-AUG-2000	2000US-0225570
PR	14-AUG-2000	2000US-0225577
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226759
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226688
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227809
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229309
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231342
PR	08-SEP-2000	2000US-0231343
PR	08-SEP-2000	2000US-0231344
PR	08-SEP-2000	2000US-0231345
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0233168
PR	14-SEP-2000	2000US-0233297
PR	14-SEP-2000	2000US-0233298
PR	14-SEP-2000	2000US-0233299
PR	14-SEP-2000	2000US-0234201
PR	14-SEP-2000	2000US-0234201
PR	14-SEP-2000	2000US-0233063
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234223
PR	25-SEP-2000	2000US-0234997
PR	25-SEP-2000	2000US-0234998
PR	25-SEP-2000	2000US-0235484
PR	27-SEP-2000	2000US-0235634
PR	27-SEP-2000	2000US-0235636
PR	29-SEP-2000	2000US-0236527
PR	29-SEP-2000	2000US-0236527
PR	29-SEP-2000	2000US-0236568
PR	29-SEP-2000	2000US-0236569
PR	29-SEP-2000	2000US-0236570
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239353
PR	13-OCT-2000	2000US-0239357
PR	20-OCT-2000	2000US-0240960
PR	20-OCT-2000	2000US-0241221
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241807
PR	20-OCT-2000	2000US-0241808

PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-024617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249219.
PR	17-NOV-2000;	2000US-0249224.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249269.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250191.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251038.
PR	05-DEC-2000;	2000US-0251188.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251899.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-451937/48.	
DR		
XX		
XX		
PT	Isolated polypeptide for treating, preventing and/or prognosing	
PT	disorders related to the musculoskeletal system including	
PT	musculoskeletal cancers and also for testing and detection e.g.	
PT	diagnosis -	
PS		
XX	Example 2; SEQ ID NO 2805; 781bp + Sequence Listing; English.	
CC		
XX	The invention relates to novel genes (AALJ3669-AALJ3666) and proteins	
CC	(AAB03087-AAB04109) associated with the musculoskeletal system useful	
CC	for preventing, treating or ameliorating medical conditions e.g. by	
CC	protein or gene therapy. The genes are isolated from a range of human	
CC	tissues disclosed in the specification. The nucleic acids, proteins,	
CC	antibodies and (anti)agonists are useful in the diagnosis, treatment	
CC	and prevention of: (a) cancer, e.g. breast and ovarian cancer and	
CC	other cancers of the adrenal gland, bone, bone marrow, breast,	

CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX

XX Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;

Query Match 54.2%; Score 880; DB 22; Length 19125;
Best Local Similarity 100.0%; Pred. No. 2.6e-217;

Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCCATCAACCTGATCGTCAGACATCCAGGACATTTGAAATGGTACATCTGCAAA 804
Db 5997 GTTGCCATCAACCTGATCGTCAGACATCCAGGACATTTGAAATGGTACATCTGCAAA 6056
QY 805 TGGCAACGAGGAGGGTCCATGGGCGAGCTAACAGCGACCTTTTCTGAGCCAGGGGAC 864
Db 6057 TGGCAACGAGGAGGGTCCATGGGCGAGCTAACAGCGACCTTTTCTGAGCCAGGGGAC 6116
QY 865 CACCTGGGATGTGACCTCTGGCAAAAGGTCACATTTGGAGTGCAGAGACGCCAC 924
Db 6117 CACCTGGGATGTGACCTCTGGCAAAAGGTCACATTTGGAGTGCAGAGACGCCAC 6176
QY 925 TGAAGGGGTGCGAGGCTTCAGGGGAGGTCTCCGCGCCGCGAGTGTGTTTACAGGACTGAG 984
Db 6177 TGAAGGGGTGCGAGGCTTCAGGGGAGGTCTCCGCGCGAGTGTGTTTACAGGACTGAG 6236
QY 985 CCTGGGAGCGCCACCCACACCTGCTTCTCTCGGCGACCCAGGGGAGTGTAGC 1044
Db 6237 CCTGGGAGCGCCACCCACACCTGCTTCTCTCGGCGACCCAGGGGAGTGTAGC 6296
QY 1045 AGCGAGGCTTCTCTCAGTCAAGAGTGAATCTCAGATGTGCTCAGACTCAACTTGGCT 1104
Db 6297 AGCGAGGCTTCTCTCAGTCAAGAGTGAATCTCAGATGTGCTCAGACTCAACTTGGCT 6356
QY 1105 GGGACACTGACAGGCGTTCTCGAGGTTTTCAGCCACTTAGGCTGTTCCGGTTTAAAGAT 1164
Db 6357 GGGACACTGACAGGCGTTCTCGAGGTTTTCAGCCACTTAGGCTGTTCCGGTTTAAAGAT 6416
QY 1165 CCTCTTAGTCTCAGGAAATGCCACAGATGTGAGAGAGCCCTGGAGGCTTGTGAG 1224
Db 6417 CCTCTTAGTCTCAGGAAATGCCACAGATGTGAGAGAGCCCTGGAGGCTTGTGAG 6476
QY 1225 GAAATGTGAGGACATTTATGGGAAATTTAGAGACAGCCCTAGACCTGGCTGGCTGAT 1284
Db 6477 GAAATGTGAGGACATTTATGGGAAATTTAGAGACAGCCCTAGACCTGGCTGGCTGAT 6536
QY 1285 GTTTTGTGACAGTAAACCCACAGTGGAGAGAGTTTTTTCAGTCTGATCTGGTTCTTA 1344
Db 6537 GTTTTGTGACAGTAAACCCACAGTGGAGAGAGTTTTTTCAGTCTGATCTGGTTCTTA 6596
QY 1345 CACACTCAGACACATACTAAAAGTTTGTGAACAAGTACTTTCCTTTTACATGTTA 1404
Db 6597 CACACTCAGACACATACTAAAAGTTTGTGAACAAGTACTTTCCTTTTACATGTTA 6656
QY 1405 CATGTCCATGATTTTCTGTTTCTGTTTCTAACAAGGCTGGTGTGGCCCTACAAAC 1464
Db 6657 CATGTCCATGATTTTCTGTTTCTGTTTCTAACAAGGCTGGTGTGGCCCTACAAAC 6716
QY 1465 CTAATTTTCATGACCCAGTGTGTTGACGTCAGCGTGGCTTACACGATATGGGAGCCAC 1524
Db 6717 CTAATTTTCATGACCCAGTGTGTTGACGTCAGCGTGGCTTACACGATATGGGAGCCAC 6776
QY 1525 TGAAGGATGTTTTCCCTTGTGCTTGTGCTTAAAGGACAGAAAGGAGGCGATGCCCT 1584
Db 6777 TGAAGGATGTTTTCCCTTGTGCTTGTGCTTAAAGGACAGAAAGGAGGCGATGCCCT 6836

QY 1585 GGAAGCACCCAGCATGACACCCAGGGCTGTGGGGGCCAG 1624
Db 6837 GGAAGCACCCAGCATGACACCCAGGGCTGTGGGGGCCAG 6876

RESULT 9

ABX59428
ID ABX59428 standard; cDNA; 19125 BP.

AC ABX59428;

DT 26-FEB-2003 (first entry)

XX cDNA encoding novel human musculoskeletal system antigen #1772.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;
XX periodontal regeneration; tissue transport; bone graft; skin aging;
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
XX cell growth; organ transplant; cell differentiation; body height;
XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
XX pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
XX depression; tendency for violence; pain; reproductive capability;
XX hormone level; endocrine level; appetite; libido; memory; stress;
XX storage capability; fat content; lipid content; protein content;
XX carbohydrate content; vitamin content; cofactor content;
XX nutritional component.

OS Homo sapiens.

XX US2002147140-A1.

PN 10-OCT-2002.

PD 17-JAN-2001; 2001US-0764877.

XX 31-JAN-2000; 2000US-179065P.
XX 04-FEB-2000; 2000US-180628P.
XX 28-JUN-2000; 2000US-214886P.
XX 07-JUL-2000; 2000US-216647P.
XX 07-JUL-2000; 2000US-216880P.
XX 11-JUL-2000; 2000US-217487P.
XX 11-JUL-2000; 2000US-217496P.
XX 14-JUL-2000; 2000US-218290P.
XX 26-JUL-2000; 2000US-220963P.
XX 26-JUL-2000; 2000US-220964P.
XX 14-AUG-2000; 2000US-224518P.
XX 14-AUG-2000; 2000US-224519P.
XX 14-AUG-2000; 2000US-225267P.
XX 14-AUG-2000; 2000US-225268P.
XX 14-AUG-2000; 2000US-225270P.
XX 14-AUG-2000; 2000US-225447P.
XX 14-AUG-2000; 2000US-225757P.
XX 14-AUG-2000; 2000US-225758P.
XX 22-AUG-2000; 2000US-226868P.
XX 30-AUG-2000; 2000US-228924P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229343P.
XX 01-SEP-2000; 2000US-229344P.
XX 01-SEP-2000; 2000US-229345P.
XX 05-SEP-2000; 2000US-229509P.
XX 08-SEP-2000; 2000US-229513P.
XX 21-SEP-2000; 2000US-234223P.
XX 21-SEP-2000; 2000US-234274P.
XX 25-SEP-2000; 2000US-234997P.
XX 27-SEP-2000; 2000US-235834P.
XX 29-SEP-2000; 2000US-236327P.

PR	29-SEP-2000;	2000US-236367P.
PR	29-SEP-2000;	2000US-236368P.
PR	29-SEP-2000;	2000US-236369P.
PR	29-SEP-2000;	2000US-236370P.
PR	02-OCT-2000;	2000US-236802P.
PR	02-OCT-2000;	2000US-237037P.
PR	02-OCT-2000;	2000US-237038P.
PR	02-OCT-2000;	2000US-237039P.
PR	02-OCT-2000;	2000US-237040P.
PR	13-OCT-2000;	2000US-239935P.
PR	20-OCT-2000;	2000US-240960P.
PR	20-OCT-2000;	2000US-241785P.
PR	20-OCT-2000;	2000US-241809P.
PR	01-NOV-2000;	2000US-244617P.
PR	17-NOV-2000;	2000US-249299P.
PR	08-DEC-2000;	2000US-251856P.
PR	08-DEC-2000;	2000US-251868P.
PR	08-DEC-2000;	2000US-251869P.
XX		
PA	(ROSE/)	ROSEN C A.
PA	(RUBE/)	RUBEN S M.
PA	(BARA/)	BARASH S C.
PI	Rosen CA, Ruben SM, Barash SC;	
DR	WPI: 2003-128199/12.	
PT	Isolated nucleic acid molecules encoding musculoskeletal system	
XX	associated polypeptides, useful for detecting disorders, e.g. cancer -	
PS	Disclosure; SEQ ID NO 2805; 321pp; English.	
XX		
CC	The invention describes an isolated nucleic acid molecule comprising a	
CC	sequence encoding musculoskeletal system associated polypeptides useful	
CC	for detecting disorders, e.g., cancer or cancer metastases, in animals	
CC	or humans. The nucleic acid: stimulates re-vascularisation of ischaemic	
CC	tissues associated with conditions such as thrombosis, arteriosclerosis,	
CC	and other cardiovascular conditions; treats wounds due to injuries,	
CC	burns, post-operative tissue repair, and ulcers; stimulates angiogenesis	
CC	and limb regeneration; stimulates neuronal growth; can treat and prevent	
CC	neuronal damage occurring in certain disorders or neurodegenerative	
CC	conditions, such as, Alzheimer's disease, Parkinson's disease, and	
CC	AIDS-related complex; stimulates chondrocyte growth, thus they can be	
CC	used to enhance bone and periodontal regeneration and aid in tissue	
CC	transports or bone grafts; prevents skin aging due to sunburn by	
CC	stimulating keratinocyte growth; prevents hair loss, since FGF family	
CC	members activate hair-forming cells and promotes melanocyte growth;	
CC	stimulates growth and differentiation of hematopoietic cells and bone	
CC	marrow cells when used in combination with other cytokines; maintains	
CC	organs before transplantation or for supporting cell culture of primary	
CC	tissues; induces tissue of mesodermal origin to differentiate in early	
CC	embryos; increases or decreases the differentiation or proliferation of	
CC	embryonic stem cells, besides, haematopoietic lineage; modulates	
CC	mammalian characteristics, such as, body height, weight, hair colour, eye	
CC	colour, skin, percentage of adipose tissue, pigmentation, size, and shape	
CC	(e.g., cosmetic surgery); modulates mammalian metabolism; changes	
CC	mammalian metal state or physical state by influencing biorythms,	
CC	cardiac, rhythms, depression, tendency for violence, tolerance for pain,	
CC	reproductive capabilities; hormonal or endocrine levels, appetite,	
CC	libido, memory, or stress; increases or decreases storage capabilities,	
CC	fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors	
CC	or other nutritional components. This sequence encodes a novel human	
CC	musculoskeletal system antigen.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from the US patent office at	
CC	ftp.seqdata.uspto.gov/sequence.html?docID=20020147140.	

```

SQ      Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
Query Match          54.2%; Score 880; DB 25; Length 19125;
Best Local Similarity 100.0%; Pred. No. 2.6e-217;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

OY	745	GTTCGCATCAACCTGATTCGTGACACATATCCAGACATTTCTGAATGCTGACATCTGCAAA	804
Db	5997	GTTCGCATCAACCTGATTCGTGACACATATCCAGACATTTCTGAATGCTGACATCTGCAAA	6056
OY	805	TGGCACCAGAGAGGGGTCCAAATGGGCGAGCTTACAAGGGGACCTTTTCTGAGCCAGGGGAC	864
Db	6057	TGGCACCAGAGAGGGGTCCAAATGGGCGAGCTTACAAGGGGACCTTTTCTGAGCCAGGGGAC	6116
OY	865	CACCCTGGGATGCTGACCTCTGGCAAA CGGTCA CATTGGAGTCCAGACAGACCCAC	924
Db	6117	CACCCTGGGATGCTGACCTCTGGCAAA CGGTCA CATTGGAGTCCAGACAGACCCAC	6176
OY	925	TGAGGGGCTGCTCGAGGCTTCAGGGATAGGTCTTCGCCGCCGATGTGTCTTCAAGGACTGAG	984
Db	6177	TGAGGGGCTGCTCGAGGCTTCAGGGATAGGTCTTCGCCGCCGATGTGTCTTCAAGGACTGAG	6236
OY	985	CCTGGGGACGCGCCACCCACACCCACTGCTCTCTCGGGGCA CCCCAGGGGAGTGTAGC	1044
Db	6237	CCTGGGGACGCGCCACCCACACCCACTGCTCTCTCGGGGCA CCCCAGGGGAGTGTAGC	6296
OY	1045	AGCAGGCGCTTCTCTCACTCAGAGTGTGAAA CTCAGATGTGTCACTCAGACTCAACTTGTCT	1104
Db	6297	AGCAGGCGCTTCTCTCACTCAGAGTGTGAAA CTCAGATGTGTCACTCAGACTCAACTTGTCT	6356
OY	1105	GGGACACTGACAGGCGCTTCTGAGGTTTTCAGCCACTTAGAGCTGTGCGGTTAAAGAT	1164
Db	6357	GGGACACTGACAGGCGCTTCTGAGGTTTTCAGCCACTTAGAGCTGTGCGGTTAAAGAT	6416
OY	1165	CCCTCTAGGTCACTGAGAAATGCCCACAGAAATGTGTCAGAAAGCTGGGAGGCTTCTGAG	1224
Db	6417	CCCTCTAGGTCACTGAGAAATGCCCACAGAAATGTGTCAGAAAGCTGGGAGGCTTCTGAG	6476
OY	1225	GAATGTGAGGCACTTATTTG3GGAATTTGAGAGACAGCCTAGACACTGTGCTGCTGAT	1284
Db	6477	GAATGTGAGGCACTTATTTG3GGAATTTGAGAGACAGCCTAGACACTGTGCTGCTGAT	6536
OY	1285	GTTTTGTGTGACAGTGAACCCACAGTGGGAGAGATTTTTCAGCTGTGATCTGCTTCTTA	1344
Db	6537	GTTTTGTGTGACAGTGAACCCACAGTGGGAGAGATTTTTCAGCTGTGATCTGCTTCTTA	6596
OY	1345	CACACTCACA CACTA ACTCAAAA GTTTGTGAAACA GATCACTTCTCTTTTACATGTTA	1404
Db	6597	CACACTCACA CACTA ACTCAAAA GTTTGTGAAACA GATCACTTCTCTTTTACATGTTA	6656
OY	1405	CATGTCTCATGTCTTCTGTTTTCTGTTTCATTA CACAAGGCTGTGTGTGGCTTACAAC	1464
Db	6657	CATGTCTCATGTCTTCTGTTTTCTGTTTCATTA CACAAGGCTGTGTGTGGCTTACAAC	6716
OY	1465	CTAATTTTCATGACCCCACTGTTGTGTGCACTCCAGGTGGCTTACACGATATGTGGAGCCAC	1524
Db	6717	CTAATTTTCATGACCCCACTGTTGTGTGCACTCCAGGTGGCTTACACGATATGTGGAGCCAC	6776
OY	1525	TGAGGAGATGTTTTCCGCCCTTGTGCTTAAAGGACAGAAAGCCAGGCGGATGCGCT	1584
Db	6777	TGAGGAGATGTTTTCCGCCCTTGTGCTTAAAGGACAGAAAGCCAGGCGGATGCGCT	6836
OY	1585	GGAAGCACCACGATCACACCTCAGGCTTGTGTGGGGGCGAG	1624
Db	6837	GGAAGCACCACGATCACACCTCAGGCTTGTGTGGGGGCGAG	6876
RESULT 10			
AAH75355 standard; cDNA; 900 BP.			
AAH75355;			
02-OCT-2001 (first entry)			
Human uridine kinase encoding cDNA.			
Human; uridine kinase; UK; ss.			


```

XX OS Homo sapiens.
XX FH Key
XX FT CDS
XX FT 8..838
XX FT /*tag= a
XX FT /product= "uridine kinase"
XX FT /note= "claimed in claim 1"
XX PN CN1287172-A.
XX PD 14-MAR-2001.
XX PF 07-SEP-1999; 99CN-0118818.
XX PR 07-SEP-1999; 99CN-0118818.
XX PA (UYFU-) UNIV FUDAN.
XX PI Yu L, Zhao Y, Zhang H;
XX WPI; 2001-409529/44.
XX P-PSDB; AAG64506.
XX Human uridine kinase and its coding sequence, preparation and
XX application.
XX PS Claim 1; Page 15(Disclosure); 20pp; Chinese.
XX CC The invention relates to human uridine kinase (UK).
XX SQ Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;

Query Match 52.6%; Score 855; DB 22; Length 900;
Best Local Similarity 97.5%; Pred. No. 2e-211;
Matches 880; Conservative 0; Mismatches 20; Indels 3; Gaps 1;

QY 87 GGCAGAGTGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCG 146
DB 1 GGCAGAGTGGCTTCGGCGGAGGCGAAG---GTGCGAGAGCCCGCGCGGAGCGAAGCCG 57
QY 147 TCCGCGACCGAGCGCCCTTCTGATAGGGGTAGCGCGCGCACTGCCAGCGGAAATCGAC 206
DB 58 TCCGCGACCGAGCGCTTCTGATAGGGGTAGCGCGCGCACTGCCAGCGGAAATCGAC 117
QY 207 CGTGTGAGAGATTCATGAGATTCCTGGGACAGAACAGAGTGGAAACAGCGGCGACCGGAA 266
DB 118 CGTGTGAGAGATTCATGAGATTCCTGGGACAGAACAGAGTGGAAACAGCGGCGACCGGTA 177
QY 267 GGTGTCATCTGAGCGCAGACAGATTCCTCAAGGTCCTGAGCGGAGAGAGGCGCAA 326
DB 178 GGTGTCATCTGAGCGCAGACAGATTCCTCAAGGTCCTGAGCGGAGAGAGGCGCAA 237
QY 327 GGCCTTGAAGAGACAGTAAATTTTGAACATCCAGATGCCCTTTGATATGATTTGATCA 386
DB 238 GGCCTTGAAGAGACAGTAAATTTTGAACATCCAGATGCCCTTTGATATGATTTGATCA 297
QY 387 CAGGACTCTGAAGAACATCGTGAAGGCGAAAAAGGTGAGGTGCGGACCTATGATTTTGT 446
DB 298 CAGGACTCTGAAGAACATCGTGAAGGCGAAAAAGGTGAGGTGCGGACCTATGATTTTGT 357
QY 447 GACACACTCAAGGTTTACAGAGACCAAGGTGTCATCCCTGCGGAGCGTGGTTCTGTTGA 506
DB 358 GACACACTCAAGGTTTACAGAGACCAAGGTGTCATCCCTGCGGAGCGTGGTTCTGTTGA 417
QY 507 GGGGATCTTGGTGTCTTACAGCCAGAGATCCGGAGATGTTCCACCTGGCCCTTTCGT 566
DB 418 GGGGATCTTGGTGTCTTACAGCCAGAGATCCGGAGATGTTCCACCTGGCCCTTTCGT 477
QY 567 GGAACACGATCTCCAGCGTCAGGCTGTCTCGAAGGTTCTCGGGAAGTGGCGGAGGAG 626
DB 478 GGAACACGATCTCCAGCGTCAGGCTGTCTCGAAGGTTCTCGGGAAGTGGCGGAGGAG 537

```

```

QY 627 GGAACCTGAGCAGATTTCTGACGAGTACACCACTTCGTGAAGCCGACCTTCGAGGAGTT 686
DB 538 GGAACCTGAGCAGATTTCTGACGAGTACACCACTTCGTGAAGCCGACCTTCGAGGAGTT 597
QY 687 CTGCTGCGCAGAAAGATATCCGATGTGATCATCCACGAGAGTGGACAAATATGTT 746
DB 598 CTGCTGCGCAGAAAGATATCCGATGTGATCATCCACGAGAGTGGACAAATATGTT 657
QY 747 TGGCATCAACCTGATGTGAGACATCCAGGACATTCGTAATGATGACATTCGCAATG 806
DB 658 TGGCATCAACCTGATGTGAGACATCCAGGACATTCGTAATGATGACATTCGCAATG 717
QY 807 GCAACCGAGAGGCTCCATAGGCGGAGCTACAGCCGACCTTTCTGAGCCAGGAGCA 866
DB 718 GCAACCGAGAGGCTCCATAGGCGGAGCTACAGCCGACCTTTCTGAGCCAGGAGCA 777
QY 867 CCTGGGATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACACCCACTG 926
DB 778 CCTGGGATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACACCCACTG 837
QY 927 AGGGGCTGCGGAGCTTCAGGCGAGTCTCCCGCCGCGCATGTGTTCAGGAGCTAGCC 986
DB 838 AGGGGCTGCGGAGCTTCAGGCGAGTCTCCCGCCGCGCATGTGTTCAGGAGCTAGCC 897
QY 987 TGG 989
DB 898 TGG 900

```

```

RESULT 11
AAK98735
ID AAK98735 standard; DNA; 834 BP.
AC AAK98735;
DT 02-MAY-2002 (first entry)
DE DNA of a human uridine kinase (UDK).
XX KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
XX KW probe; chromosome localization study; tissue expression; gene therapy;
XX KW antibody; vaccine; human ovarian cancer; immunological disorder;
XX KW human colon carcinoma; immunogen; ds.
OS Homo sapiens.
XX FH Key
XX FT CDS
XX FT 1..780
XX FT /*tag= a
XX FT /partial
XX FT /note= "No stop codon"
XX PN WO200172963-A2.
XX PD 04-OCT-2001.
XX PF 27-MAR-2001; 2001WO-US09663.
XX PR 27-MAR-2000; 2000US-0536647.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI Ho YS, Johnson RK;
XX WPI; 2001-626259/72.
XX P-PSDB; AAO14412.
XX Novel human uridine kinase polypeptides useful for treating cancers,
XX and to identify agonists and antagonists of the polypeptide useful for
XX treating conditions associated with uridine kinase imbalance.
XX Claim 5; Page 22-23; 31pp; English.

```


CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilizing uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localization studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and
 CC immunological disorders. This polynucleotide sequence represents the DNA
 CC of a human uridine kinase of the invention.

XX Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 51.3%; Score 832.4; DB 23; Length 834;
 Best Local Similarity 99.9%; Pred. No. 1.4e-205;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 153
 Db 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 60
 QY 154 CAGCGGCCCTCTCTGATAGGGGCTGAGCGGCGCACTCCAGCGGAAAGTGCACCTGTGT 213
 Db 61 CACCGGCCCTCTCTGATAGGGGCTGAGCGGCGCACTCCAGCGGAAAGTGCACCTGTGT 120
 QY 214 GAGAGATCATGAGGCTGCTGGGACAGACGAGGTGGAACAGGGGCGGAAAGTGTGTC 273
 Db 121 GAGAGATCATGAGGCTGCTGGGACAGACGAGGTGGAACAGGGGCGGAAAGTGTGTC 180
 QY 274 ATCTGAGCCAGGACAGGTTCTACAAAGTCTTGAACGCGACAGACGAAAGGCGCTTG 333
 Db 181 ATCTGAGCCAGGACAGGTTCTACAAAGTCTTGAACGCGACAGACGAAAGGCGCTTG 240
 QY 334 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGGACT 393
 Db 241 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGGACT 300
 QY 394 CTGAGAGCATCGTGGAGGGCAAAAGCGTGGAGGTGCGACCTATGATTTTGTGACACAC 453
 Db 301 CTGAGAGCATCGTGGAGGGCAAAAGCGTGGAGGTGCGACCTATGATTTTGTGACACAC 360
 QY 454 TCAAGGTTTACAGAGACCAACGCTGTCTACCTGCGGACGTGTTCTGTTGAGGGCATC 513
 Db 361 TCAAGGTTTACAGAGACCAACGCTGTCTACCTGCGGACGTGTTCTGTTGAGGGCATC 420
 QY 514 TTGGTGTTCACACCCAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGTGACACC 573
 Db 421 TTGGTGTTCACACCCAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGTGACACC 480
 QY 574 GACTCCGAGTCAGGCGTCTCGAAGAGTCTCCGGGACGTGGGCGGAGGAGGAGGACTG 633
 Db 481 GACTCCGAGTCAGGCGTCTCGAAGAGTCTCCGGGACGTGGGCGGAGGAGGAGGACTG 540
 QY 634 GACAGATTCAGCGCAGTACACCACTTGTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 693
 Db 541 GACAGATTCAGCGCAGTACACCACTTGTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 600
 QY 694 CCGAGCAAAAGATGCTCCGATGTGATATCCACAGAGAGTGAACAATATGTTGCCATC 753
 Db 601 CCGAGCAAAAGATGCTCCGATGTGATATCCACAGAGAGTGAACAATATGTTGCCATC 660
 QY 754 AACCTGATGTGAGACATCCAGGACATCTGAAATGTGATCAATCTGCAAAATGGACCGA 813

Db 661 AACCTGATGTGAGACATCCAGGACATTTGAAATGTGATCTGCAAAATGGACCGA 720
 QY 814 GAGGGGTCCAAATGAGGCGGAGCTACAAAGGACCTTTTGTGAGCCAGGAGACCACTGGG 873
 Db 721 GAGGGGTCCAAATGAGGCGGAGCTACAAAGGACCTTTTGTGAGCCAGGAGACCACTGGG 780
 QY 874 ATCTGACCTCTGGCAAAAGGTACATTTTGGAGTCCAGACGACGACCCCACTGA 927
 Db 781 ATCTGACCTCTGGCAAAAGGTACATTTTGGAGTCCAGACGACGACCCCACTGA 834

RESULT 12

AAH04832
 ID AAH04832 standard; cDNA, 753 BP.

AAH04832;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1667.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; sg.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000BP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,

Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length cDNAs -

Claim 1; SEQ ID 1667; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

to the complementary strand of a polynucleotide which comprises one of

the 5602 nucleotide sequences defined in the specification, where the

oligonucleotide comprises at least 15 nucleotides; or (b) a combination

of an oligonucleotide comprising a sequence complementary to the

complementary strand of a polynucleotide which comprises a 5'-end

sequence and an oligonucleotide comprising a sequence complementary to a

polynucleotide which comprises a 3'-end sequence, where the

oligonucleotide comprises at least 15 nucleotides and the combination of

the 5'-end sequence/3'-end sequence is selected from those defined in

the specification. The primer sets can be used in antisense therapy and

in gene therapy. The primers are useful for synthesizing polynucleotides,

particularly full-length cDNAs. The primers are also useful for the

detection and/or diagnosis of the abnormality of the proteins encoded by

the full-length cDNAs. The primers allow obtaining of the full-length

cDNAs easily without any special methods. AAH03166 to AAH13628 and

AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to

AAB95893 represent human amino acid sequences; and AAH13629 to AAH16332

represent oligonucleotides, all of which are used in the exemplification

Db 518 ACAATGAACATCTCTCAAAAACCTCAAGAAATCACTGAAGGAGAAAACAGTCCAGATCC 577
 Qy 431 CGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCGGTGGTCTACCCGTGGG 490
 Db 578 CCGGTATGACCTTTGTCTCCCATTTCCGGAAGAGAGACAGTTACTGTCTATCCCGAG 637
 Qy 491 ACGTGTCTCTGTGGAGGACATCTTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCC 550
 Db 638 ACGTGTCTCTGTGAAGGAGATCTGGCTTCTACTCCAGAGAGTACGAGACCTGTTCC 697
 Qy 551 ACGTGTCTCTGTGACACACCACTCCGACGTAGGCTGTCTGAAGAGTCTCCGGG 610
 Db 698 AGATGAACCTTTTGTGATACAGATGGGACACCCGGCTCTCAGCAGAGATATTAAAGG 757
 Qy 611 ACGTGTCTCTGTGAGGAGACCTGAGACAGATCTCAGCAGAGTCCAGTCTGTA 667
 Db 758 ACATCAGGAGAGAGAGAGGAGATCTTGAGCAGATTTTATCTCAGTACATTCAGTTCTCA 817
 Qy 668 AGCCGGCTCTGAGAGATCTGCTGCGGACAGAAAGATGCGAGTGTGATCAATCCGAC 727
 Db 818 AGCTGCTCTTGAGGAATTTGCTGCTGCGCAACAAAGATATGCTGATGATCAATCCCTA 877
 Qy 728 GAGAGTGAACAATATGTTGCTCCATCACTGATGTCAGACATCCAGACATTTCTGA 787
 Db 878 GAGGTGAGATATCTGTGGCCATCAACCTCATCTGTGACACATCCAGACATCCCTGA 937
 Qy 788 ATGTGACATCTGCAATGCA 809
 Db 938 ATGAGGGCCCTCCAAAGGCA 959

RESULT 15
 ID AA157850 standard; cDNA; 1402 BP.
 AA157850;

AC AA157850;
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 53.

KW Human; noctropic; immunosuppressant; cyrostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

OS Homo sapiens.

PN MO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM8694.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 53; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;

Qy Query Match 23.2%; Score 377; DB 22; Length 1402;

Db Best Local Similarity 69.0%; Pred. No. 2, 1e-87;

Matches 546; Conservative 0; Mismatches 240; Indels 5; Gaps 2;

Qy 22 GCGCTGGGCGGCGCGCGCGCGGAAAGGCGCGCGCGGACCCGAGTCCGCGGA 81
 Db 211 GCGGAGGAGATCCGACGCGCGCGCGCGCGCGCGTCCGTCGACAGGACGCG 270
 Qy 82 GCGGAGGCGCGAGATGCGTCCGCGGAGGAGAGATCGGAGAGCCCGCGCGAGGCC 141
 Db 271 GAGAGGCGCGCGCGGAAACATGCGCGGAGACAGACAGACCTTCAGAAACACAGCA 330
 Qy 142 GACCGTCCGACGAGCGCGCGCTTCTGATAGGCGGTGAGCGCGGACCTGCGAGCGGAA 201
 Db 331 GCCCAACGCGCGC--GAGCCCTTCTTATAGGGGTGACGCGGGAACAGTACGCGCAAG 388
 Qy 202 TCGACCGTGTGAGAAAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACGCGGAC 261
 Db 389 TCTTCGCGTGTCTAAAGATCGGACGCTCTGCGGACAGAAATGAGTGCATATCCGAC 448
 Qy 262 CGGAGGTGTCTCTGAGCGCGAGAGAGTTTACAAGTCTCTGACGCGAGCGAGAG 321
 Db 449 AAGCAGGTGTCTCTGAGCGCGAGATGCTTACCGTGTCTTACCTTCGAGCGAGAG 508
 Qy 322 GCCAAGGCTTGAAGAGACAGTCAATTTTGAACATTCAGATGCTTTGATTAATTTG 381
 Db 509 GCCAAGGCTTGAAGAGGCGAGTTCATTCAGACACCGGATGCTTTGATTAATTAATTC 568
 Qy 382 ATGCAAGAGATCTGTAAGAACATGAGAGGCAAAACGAGTGGCGGACCTATGAT 441
 Db 569 ATTCTCAAAACATCAAGAAATCACTGAAGGGAAGAAACAGTCCAGATCCCGGTATGAC 628
 Qy 442 TTTGTGACACCTCAAGTTCACAGAGACCAAGTGTCTTACCTTCGAGAGTGTTCG 501
 Db 629 TTTGTCTCCATTCGCGAAGAGAGAGACATTTACTGTCTATCCGCAACGAGTGTTC 688
 Qy 502 TTTGAGGCGATCTTGTGTTTCTACAGCGAGAGATCCGGAAGATGTTCCAGTCCGCTC 561
 Db 689 TTTGAGGAGATCTTGCGCTTCTACTCCAGAGAGTACAGAGACCTGTTCAGATGAAGCTT 748
 Qy 562 TTGCTGAGACCGGATCCGACGTCGAGTGTCTCGAAGAGTTCCTCGGAGAGTGCCTC 618
 Db 749 TTTGTGATATCAGATCCGAGACACCGGCTCTCAGCAGAGTATTAGGAGACATCAAGAG 808
 Qy 619 CGAGGAGAGGACCTGAGAGAGATTCGACGAGTACACACATTCCTGTAAGCGGCGCTTC 678
 Db 809 AGAGGAGGAGATCTTGAAGAGATTTATCTCAGTACATTACGTTGCTCAAGCCTGCTTT 868

QY	679	GAGAGTTCTGCTGCGCGACAAAGAATAGCCGATGTGATCATCCACGAGAGTGCAC	738
Db	869	GAGGATTTCTGCTTGCCACAAAGAGTATGCTGATGTGATCATCCCTAGAGGTGCAGAT	928
QY	739	AATATGTTGCCATCAACTGATCGTGACGACATCCAGGACATTTCTGAATGTTGACATC	798
Db	929	AATCTGTGGCCATCAACTCATCGTGACGACATCCAGGACATCCTGAATGAGGGGCC	988
QY	799	TGCAATGGCA	809
Db	989	TCCAAACGGCA	999

Search completed: November 24, 2003, 21:42:25
Job time : 503.186 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2003, 16:46:08 ; Search time 6310.34 Seconds

(without alignments)
10528.307 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 1624

Sequence: 1 gtcggggtcgcctccgcaccc.....ccagctctgtgaggccag 1624

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb ba: *
2: gb htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_or: *
21: em_ov: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pin: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htg_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1624	100.0	1624	6	AX449217 Sequence
2	1599.4	98.5	2160	6	BD157613 Primer fo
3	1599.4	98.5	2160	6	AK022317 Homo sapi
4	1453.4	89.5	2228	9	AK057848 Homo sapi
5	1397.4	86.0	2072	9	BC015547 Homo sapi
6	1156.8	71.2	1395	9	AF125106 Homo sapi
7	1018.8	62.7	1022	9	AF237290 Homo sapi
8	880	54.2	147492	9	AL358781 Human DNA
9	834	51.4	834	6	AX449219 Sequence
10	832.4	51.3	834	6	AF254133 Homo sapi
11	710	43.7	753	6	BD146824 Primer fo
12	666.6	41.0	1959	10	BC025146 Mus muscu
13	623	38.4	1810	10	MUSURKI
14	597.4	36.8	952	10	MUSURKI
15	594.6	36.6	25493	2	AC078885 Mus muscu
16	574	35.3	192336	2	AC118474 Mus muscu
17	574	35.3	211075	10	AC076974 Mus muscu
18	483.6	29.8	734	6	AX540411 Sequence
19	379.6	23.4	1209	9	BC002906 Homo sapi
20	379.6	23.4	1322	6	AX135546 Sequence
21	374	23.0	744	9	AB062451 Homo sapi
22	374	23.0	1121	9	AF236637 Homo sapi
23	364.6	22.5	1784	5	BC045968 dario rer
24	363.2	22.4	1312	10	BC023789 Mus muscu
25	361.6	22.3	1354	10	AF236636 Mus muscu
26	300.4	18.5	1467	3	AY119583 Drosophill
27	255.6	15.7	841	11	BV017303 S212P6359
28	182.2	11.2	831	9	D78335 Human mRNA
29	177.2	10.9	1847	6	AX463241 Sequence
30	177.2	10.9	1847	9	AK000524 Homo sapi
31	175.8	10.8	1875	9	BC033078 Homo sapi
32	175.6	10.8	1813	6	AX405842 Sequence
33	155.6	9.8	1823	10	BC016535 Mus muscu
34	155.6	9.6	156905	2	AC020006 Drosophill
35	155.6	9.6	199016	3	AC008204 Drosophill
36	155.6	9.6	206741	3	AE003747 Drosophill
37	153.2	9.4	314	10	AB030700 Rattus no
38	150.2	9.2	336	6	BT006860 Homo sapi
39	150.2	9.2	336	12	BT007548 Synthetic
40	146.2	9.0	233210	2	AC097693 Rattus no
41	146.2	9.0	244105	2	AC098897 Rattus no
42	145.2	8.9	3246	3	AY129436 Drosophill
43	140.8	8.7	241882	10	AL808027 Mouse DNA
44	140.8	8.7	254677	2	AC068494 Mus muscu
45	136.2	8.4	655	11	BV077340

ALIGNMENTS

RESULT 1
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Glucksmann, M. A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

	REFERENCE	Eukaryota; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	TITLE	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakematsu,A., Nagai,K. and Otsuki,T.
	JOURNAL	Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002191363-A 12456 09-JUL-2002;
	COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/12456 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAZUO SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10,PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N5/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers (95) . (925) . FT CDS Location/Qualifiers 1..2160 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
	FEATURES	BASE COUNT 457 a 591 c 671 g 441 t
	ORIGIN	
Qy	Query Match	98.5%; Score 1599.4; DB 6; Length 2160;
Db	Best Local Similarity	99.7%; Pred. No. 0;
	Matches 1614; Conservative	; Mismatches 1; Indels 4; Gaps 1;
Qy	6	GTCGCTCCGACCTCGGGCTGGGGGGCGCGCGCGCGCGGAAGGGCGGGCGG 65
Db	7	GTCGCTCCGACCTCGGGCTGGGGGGCGCGCGCGCGCGCGCGCGCGCGGG 66
Qy	66	GACCCGATGCCGGGAGCGGAGCCGGAATGGCTTCGGCGGAGCGCAAGACTCGAGAAG 125
Db	67	GACCCGATGCCGGGAGCGGAGCCGGAATGGCTTCGGCGGAGCGCAAGACTCGAGAAG 126
Qy	126	CCCCGC GCCGAGGCCGACCGTCCGACCAAGCGCCCTTCTCATAGAGGTGAGCGCGG 185
Db	127	CCCCGC GCCGAGGCCGACCGTCCGACCAAGCGCCCTTCTCATAGAGGTGAGCGCGG 186
Qy	186	CACATGCCAGCGGAGTAGTCACTGTGTGAGAAATCATTGAGTTGCTGGACAAGAGA 245
Db	187	CACATGCCAGCGGAGTAGTCACTGTGTGAGAAATCATTGAGTTGCTGGACAAGAGA 246
Qy	246	GGTGAACAGCGGCGAGGAGGTGTATCTGTGACGACGAGCAGATTCTAACAGTCT 305
Db	247	GGTGAACAGCGGCGAGGAGGTGTATCTGTGACGAGCAGAGGTTCTAACAGTCT 306
Qy	306	GACGCGAGCAGAGGCGCAAAGCCCTTGAAGAGACGTACATTTTGACCATCCAGATGC 365
Db	307	GACGCGAGCAGAGGCGCAAAGCCCTTGAAGAGACGTACATTTTGACCATCCAGATGC 366
Qy	366	CTTTGATATATTTGATGCAAGGACTCTGAAGAACTCGTGAGGGCGAAAACGTTGA 425
Db	367	CTTTGATATATTTGATGCAAGGACTCTGAAGAACTCGTGAGGGCGAAAACGTTGA 426
Qy	426	GGTGCGGACCATGTTTTTGACACATCAAGATTACAGAGACACAGGTGTATACC 485
Db	427	GGTGCGGACCATGTTTTTGACACATCAAGATTACAGAGACACAGGTGTATACC 486
Qy	486	TGCGAGCGTGTCTGTGTGAGGGCATTTGGTGTCTTACAGCCAGAGATCCGGACAT 545
Db	487	TGCGAGCGTGTCTGTGTGAGGGCATTTGGTGTCTTACAGCCAGAGATCCGGACAT 546
Qy	546	GTTCCACCTGGGCTCTTGTGTGACACCGACTCCGAGGTCAAGCTGTCTCGAAGTTCT 605

Dp	547	GTTCACACTGGCGCTCTTTCGATGGAACCCGACTCCGAGGTCAAGGCTGTCTGAAAGATTCT	606
OY	606	CCGGGAGCTGCGCCGAGAGGAGGACCTTGAGACAGATTCTGACGAGTACACCACTTCTGT	665
Dp	607	CCGGGAGCTGCGCCGAGAGGAGGACCTTGAGACAGATTCTGACGAGTACACCACTTCTGT	666
OY	666	GAACCCGGGCTTCCAGGAGTTCTGCTCCGAGCAAAAGAAATCCGATGATCATCTCC	725
Dp	667	GAACCCGGGCTTCCAGGAGTTCTGCTCCGAGCAAAAGAAATCCGATGATCATCTCC	726
OY	726	ACGAGAGATGGACAAATATGTTGGCATCAACCTGATCGTCAGACATCCAGACATTCT	785
Dp	727	ACGAGAGATGGACAAATATGTTGGCATCAACCTGATCGTCAGACATCCAGACATTCT	786
OY	786	GAATGTGACATCTGCAAAATGGCACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGGAC	845
Dp	787	GAATGTGACATCTGCAAAATGGCACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGGAC	846
OY	846	CTTTTCTGAGCCAGGGGACCACTCTGGGATGCTGACTCTGGGCAAAAGGTCACTTTTGA	905
Dp	847	CTTTTCTGAGCCAGGGGACCACTCTGGGATGCTGACTCTGGGCAAAAGGTCACTTTTGA	906
OY	906	GTCCAGACGACGACCCCACTGAGGGGCTCCGAGACTCAAGGACAGGTCTCCGCCCGGACA	965
Dp	907	GTCCAGACGACGACCCCACTGAGGGGCTCCGAGACTCAAGGACAGGTCTCCGCCCGGACA	966
OY	966	TGTGTGTTCAAGGAGCTGAGCCTGGGGGACGCCCAACCACTGCTTCTCTCCGCGC	1025
Dp	967	TGTGTGTTCAAGGAGCTGAGCCTGGGGGACGCCCAACCACTGCTTCTCTCCGCGC	1026
OY	1026	ACCCACGAGGAGGTGTAGCAGCGAGGCTTCTCACTCAGAGATGGAACTCAAGTGTGT	1085
Dp	1027	ACCCACGAGGAGGTGTAGCAGCGAGGCTTCTCACTCAGAGATGGAACTCAAGTGTGT	1086
OY	1086	CACATCAGACTCAATCTGCTGGGACACTGACAGGGGTTCTGAGGTTTTCACCCACTTAGG	1145
Dp	1087	CACATCAGACTCAATCTGCTGGGACACTGACAGGGGTTCTGAGGTTTTCACCCACTTAGG	1146
OY	1146	CTCGTTGCGGTTTAAAGATCCCTCTAGATCACTGAGAAATGCCCCAGAAATGTCAGAGAG	1205
Dp	1147	CTCGTTGCGGTTTAAAGATCCCTCTAGATCACTGAGAAATGCCCCAGAAATGTCAGAGAG	1206
OY	1206	CTTGGAGAGGCTTCTGTGAGGAAATGTGAGGCACTTATTTGGGGAATTTGAGGACAGCCT	1265
Dp	1207	CTTGGAGAGGCTTCTGTGAGGAAATGTGAGGCACTTATTTGGGGAATTTGAGGAGCAGCCT	1266
OY	1266	AGACACTGCGCTGCGCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGAGTTTTTTC	1325
Dp	1267	AGACACTGCGCTGCGCTGATGTTTGTGTGACAGTGAACCCACAGTGGGAGAGAGTTTTTTC	1326
OY	1326	CAGTCTGATCTGTTCTTACACACTCACACATPACTCAAAAGTTTTGTGAACAAGTAC	1385
Dp	1327	CAGTCTGATCTGTTCTTACACACTCACACATPACTCAAAAGTTTTGTGAACAAGTAC	1386
OY	1386	TTTTCCTTTTTCATAGTTACATGTCATGTCATGTTTCTGTTTCTGTTTCATPACACAAG	1445
Dp	1387	TTTTCCTTTTTCATAGTTACATGTCATGTCATGTTTCTGTTTCTGTTTCATPACACAAG	1446
OY	1446	CTGTTGTGGCTTACAAACTTAATTTATGACCCAGTGGTTTGTGAGTCCAGCTGGCCTA	1505
Dp	1447	CTGTTGTGGCTTACAAACTTAATTTATGACCCAGTGGTTTGTGAGTCCAGCTGGCCTA	1506
OY	1506	CACGGATATGAGGGGAGCACCTGAGGAGATGTTTTCGCCCTTGCTTGTGACCTTAAAGGACGA	1565
Dp	1507	CACGGATATGAGGGGAGCACCTGAGGAGATGTTTTCGCCCTTGCTTGTGACCTTAAAGGACGA	1566
OY	1566	GAAGCGAGGCGATGCCCTGGAAGCACCCAGCATCACACCGAGCTTGTGCGGGGCCAG	1624
Dp	1567	GAAGCGAGGCGATGCCCTGGAAGCACCCAGCATCACACCGAGCTTGTGCGGGGCCAG	1625

LOCUS	AK022317	2160 bp	mRNA	linear	PRI 01-AUG-2002
DEFINITION	Homo sapiens cDNA FLJ12255 f18, clone MAMMA1001476, highly similar to URIDINE KINASE (BC 2.7.1.48).				
ACCESSION	AK022317				
VERSION	AK022317.1	GI:10433687			
KEYWORDS	oligo capping; f18 (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 Isogai,T., Oca,T., Hayashi,K., Sugiyama,T., Osuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shitatori,A., Sudo,H., Wagatsuma,M., Hosoihi,T., Kakui,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Satou,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K., Masuno,Y., Ninomiya,K. and Iwayanagi,T.				
TITLE	NEDO human cDNA sequencing project				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 2160)				
AUTHORS	Isogai,T. and Otsuki,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genom@shri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'-& 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.				
COMMENT	Location/Qualifiers				
FEATURES	1..2160				
SOURCE	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="MAMMA1001476"				
	/tissue_type="Mammary gland"				
	/clone_id="MAMMA1"				
CDS	/note="Cloning vector: pMEISFL3"				
	95..928				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="BAB14010.1"				
	/db_xref="GI:10433688"				
	/translation="MASAGDECCSPAPADRPFRPPLIVSGSTASGKTEVKIM ELUNGNEVQRKVALLSDGFYVLTAQCAKALKQYNVDHPDAIDNDLMRTLL NIVEKTALVEPDVRFTSHRLPETTVVPADVLFEGILVFSOEIRDMFLRFVTDI DSDVALSRRLVDYRGRDLLEQLITQYTFPKPAFEFCLEPTKYADVIIRGVDMVM AIIINVGHIDILINDGICKMRHGSGNSYKRTPESPDDHPOMLTSGKSHLESSRR H"				
BASE COUNT	457 a 591 c 671 g 441 t				
ORIGIN					
Query Match	98.5%; Score 1599.4; DB 9; Length 2160;				
Beat Local Similarity	99.7%; Pred. No. 0;				
Matches 1614; Conservative	0; Mismatches 1; Indels 4; Gaps 1;				
Oy	6 GTGCCTCCGACTTCGGCGCTTGAGGCGGCGGCCGGGCCCCGGGAAGAGGCGGGCGGGG 65				
Dd	7 GTTGCCTCCGACTTCGGCGCTTGAGGCGGCGGCCGGGCTTGAGGAAGAGGCGGGCGGG 66				
Oy	66 GATCCCATATGCGGGGAGCGGAGCCGAGAATTGGCTTCGGCGGGAGCGGAAGACTGGAAAG 125				
Dd	67 GATCCGATGCGGGGAGCGGAGCCGAGAATGCTTCGCGGGAGGCGGAAGACTGGGAAG 126				
Oy	126 CCCCAGCGCGAGAGCGCAGACCGTCCGACAACAAGGCGGCGCTTCGTATAGAGGATGAGCGGCG 185				
Dd	127 CCCCGCGCGAGAGCGCAGACCGTCCGACAACAAGGCGGCGCTTCGTATAGAGGATGAGCGGCG 186				
Oy	186 CACTCCAGCGGGAAGTCGACCCTGTGTGAAGAATCATGAGTTCTGGGACAGAACGA 245				

Db	187	CACGCCAGCGGGAAGTGCACCGTGTGTGAGAGATATGGAAGTTGCTGGGACAAGACGA	246
Qy	246	GGTGAACAGCGGAGCGGGAAGGTGTATCTTGAAGCCAGAGA CAGTTCTACAAGGTCCT	305
Db	247	GGTGAACAGCGGAGCGGGAAGGTGTATCTTGAAGCCAGAGA CAGTTCTACAAGGTCCT	306
Qy	306	GACGGCAGAGCAGAAAGGCCAAGGCTTTGAAAAGCAGTACAAATTTTGA CCAATCCAGATGC	365
Db	307	GACGGCAGAGCAGAAAGGCCAAGGCTTTGAAAAGCAGTACAAATTTTGA CCAATCCAGATGC	366
Qy	366	CTTTGATTAATGATTGATGCA CAGACCTCTGAAGAA CAGTGTGGAGGGCAAAA CCGGTGA	425
Db	367	CTTTGATTAATGATTGATGCA CAGACCTCTGAAGAA CAGTGTGGAGGGCAAAA CCGGTGA	426
Qy	426	GGTGCAGACCTATGATTTTGTGACA CACTAGGTTTCCAGAGACCA CCGTGTGTCTACCC	485
Db	427	GGTGCAGACCTATGATTTTGTGACA CACTAGGTTTCCAGAGACCA CCGTGTGTCTACCC	486
Qy	486	TGCGGACGTGTTCTGTTTGAAGGGCATCTGTGTCTTCA CAGCCAGAGATCCGGGACAT	545
Db	487	TGCGGACGTGTTCTGTTTGAAGGGCATCTGTGTCTTCA CAGCCAGAGATCCGGGACAT	546
Qy	546	GTTCCACCTGGGCTCTTTCGTGGA CACCGA CTTCCGAGGTGCTGTGGAAGATCT	605
Db	547	GTTCCACCTGGGCTCTTTCGTGGA CACCGA CTTCCGAGGTGCTGTGGAAGATCT	606
Qy	606	CCGGGACGTGGCCGCGAGGAGGAGCCTGTGAGCAGATTCTGACG CAGTACACCA CTTTCGT	665
Db	607	CCGGGACGTGGCCGCGAGGAGGAGCCTGTGAGCAGATTCTGACG CAGTACACCA CTTTCGT	666
Qy	666	GAAAGCCGGCTTTCGAGGAGTTCTGCTCTGCGGACA AAGAGTATGCCAGTGTATCATCCC	725
Db	667	GAAAGCCGGCTTTCGAGGAGTTCTGCTCTGCGGACA AAGAGTATGCCAGTGTATCATCCC	726
Qy	726	ACGAGAGATGGA CAAATATGTTGTCATCA CACTGATGTCAGACATCCAGACATTTCT	785
Db	727	ACGAGAGATGGA CAAATATGTTGTCATCA CACTGATGTCAGACATCCAGACATTTCT	786
Qy	786	GAATGTTGACATCTGCAAAATGCA CACGAGGAGGTTCCAAATGGGCGAGCTTACAAGCCGAC	845
Db	787	GAATGTTGACATCTGCAAAATGCA CACGAGGAGGTTCCAAATGGGCGAGCTTACAAGCCGAC	846
Qy	846	CTTTTCTGAGCCAGGGGACCA CCGTGGAGTGTGACCTGTGGCAAA CCGTCA CATTGGGA	905
Db	847	CTTTTCTGAGCCAGGGGACCA CCGTGGAGTGTGACCTGTGGCAAA CCGTCA CATTGGGA	906
Qy	906	GTTCCAGCAGCA CCCCACCTGAGGGGCTGCCGAGCTCAGGGCAGAGTCTCCCCCGCGCA	965
Db	907	GTTCCAGCAGCA CCCCACCTGAGGGGCTGCCGAGCTCAGGGCAGAGTCTCCCCCGCGCA	966
Qy	966	TGTGTGTTCA GAGACTGAGCTGTGGGAGCGCCACCCACCA CCACTGCTTCCCTCCGCGCGC	1025
Db	967	TGTGTGTTCA GAGACTGAGCTGTGGGAGCGCCACCCACCA CCACTGCTTCCCTCCGCGCGC	1026
Qy	1026	ACCCAGGGGAGTGTAGCAGCGGCTTCTCACTCAGAGTGA AACTCAGATGTGT	1085
Db	1027	ACCCAGGGGAGTGTAGCAGCGGCTTCTCACTCAGAGTGA AACTCAGATGTGT	1086
Qy	1086	CACCTCAGACTCA ACTTGTGTGGGACA CTTGACAGGGGTTCTTGAAGTTTTCAGCCACTTAA G	1145
Db	1087	CACCTCAGACTCA ACTTGTGTGGGACA CTTGACAGGGGTTCTTGAAGTTTTCAGCCACTTAA G	1146
Qy	1146	CTCGTGTGGTTTAAAGATCCCTCTAGAGTCA CTTAGAAAATGCCA CAGAAATGTGCGAGAG	1205
Db	1147	CTCGTGTGGTTTAAAGATCCCTCTAGAGTCA CTTAGAAAATGCCA CAGAAATGTGCGAGAG	1206
Qy	1206	CTTGGAGGCTTCTGTGAGGATGTGAGGCACATTAATGGGGAATTTGAGAGACAGCCT	1265
Db	1207	CTTGGAGGCTTCTGTGAGGATGTGAGGCACATTAATGGGGAATTTGAGAGACAGCCT	1266
Qy	1266	AGACACTGGCTGAGTGTGTTTGTGACAGTGA ACCCAGTGGGAGAGAGTTTTC	1325

Db 1267 AGACACTGCTGGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGATTTTTC 1326
 QY 1326 CAGTGTGATCTGTTCTTACACTCACAACATTAATCAAAAGTTTGTGAACAAGTAC 1385
 Db 1327 CAGTGTGATCTGTTCTTACACTCACAACATTAATCAAAAGTTTGTGAACAAGTAC 1382
 QY 1386 TTTCCCTTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAACAACAAG 1445
 Db 1383 TTTCCCTTTTACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAACAACAAG 1442
 QY 1446 CTGTTGTGCTTACAAACCTAATTTATGACCCAGTGTGTGACGTCCAGCTGCTTA 1505
 Db 1443 CTGTTGTGCTTACAAACCTAATTTATGACCCAGTGTGTGACGTCCAGCTGCTTA 1502
 QY 1506 CAGGATATGCGGAGCACTGAGGAGTGTTCCTCCCTGCTTGTGCTTAAAGGAGA 1565
 Db 1503 CAGGATATGCGGAGCACTGAGGAGTGTTCCTCCCTGCTTGTGCTTAAAGGAGA 1562
 QY 1566 GAAGCGAGCGGAGTGCCTGGAAGCACCCAGCATCAACCCAGGCTGTGCGGGGCGAG 1624
 Db 1563 GAAGCGAGCGGAGTGCCTGGAAGCACCCAGCATCAACCCAGGCTGTGCGGGGCGAG 1621
 RESULT 4
 AK057848 2228 bp mRNA linear PRI 31-OCT-2001
 LOCUS DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
 AK057848
 ACCESSION AK057848.1 GI:16553809
 VERSION oligo capping; fis (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Buxatoyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
 Fukunumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H.,
 Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T.,
 Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A.,
 Kawakami, B., Nagai, K., Isegai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2228)
 Sugano, S. and Suzuki, Y.
 Direct Submission
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology (RAB); cDNA library
 construction and 5'-end one pass sequencing; Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing; RAB; clone selection for
 full insert sequencing; RAB and Helix Research Institute.
 FEATURES
 source
 1. 2228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBR05878"
 /issue_type="brain"
 /clone_id="CBR"
 /note="cloning vector: pMB18SFL3"
 BASE COUNT 478 a 598 c 696 g 456 t
 ORIGIN
 Query Match 89.5%; Score 1453.4; DB 9; Length 2228;
 Best Local Similarity 93.7%; Pred. No. 1.9e-296;
 Matches 1589; Conservative 0; Mismatches 1; Indels 105; Gaps 3;

QY 31 GGGCGCGCGCGCGCGCGGGAAGGGGCGGGGCGGGGCGCGGACCCGATGCGCGGGAAGCGGCGC 90
 Db 1 GGGCGCGCGCGCGCGCGGGAAGGGGCGGGGCGGGGCGCGGACCCGATGCGCGGGAAGCGGCGC 60
 QY 91 GAGATGGCTTCGCGCGGGAAGCGAAGACTGCGAGAGCCCCGCGCGGAGCGGACCTGCC 150
 Db 61 GAGATGGCTTCGCGCGGGAAGCGAAGACTGCGAGAGCCCCGCGCGGAGCGGACCTGCC 120
 QY 151 CACACGCGCCCTTCTGATAGGGGTGAGCGGGGCGCATGCGACGCGGAAG----- 201
 Db 121 CACACGCGCCCTTCTGATAGGGGTGAGCGGGGCGCATGCGACGCGGAAGCGGACGCGG 180
 QY 202 -----TGCACGCTGTGGAAGATCATATGAG 228
 Db 181 TTCCAGCGGGGAATCCCGCTTCTGTGTCTTCAGTCCACCTGTGTGGAAGATCATATGAG 240
 QY 229 TTGCTGGACAGAACGAGGTGGAACAGCGGACGCGGAGAGTGTCTCTGAGCCGAGAC 288
 Db 241 TTGCTGGACAGAACGAGGTGGAACAGCGGACGCGGAGAGTGTCTCTGAGCCGAGAC 300
 QY 289 AGGTTCTACAGGTCTGACGCGCAGCAGAAAGGCCAAGCCTTGAAAGGACATCAAT 348
 Db 301 AGGTTCTACAGGTCTGACGCGCAGCAGAAAGGCCAAGCCTTGAAAGGACATCAAT 360
 QY 349 TTGACCATCCAGATGCTTTGATATGATTTGATGACACAGACCTCTGAAGACATCTCG 408
 Db 361 TTGACCATCCAGATGCTTTGATATGATTTGATGACACAGACCTCTGAAGACATCTCG 420
 QY 409 GAGGCGAAACGCTGAGGTGCGGACCTTGTGATTTTGTGACACACTCAAGTTTACAGAG 468
 Db 421 GAGGCGAAACGCTGAGGTGCGGACCTTGTGATTTTGTGACACACTCAAGTTTACAGAG 480
 QY 469 ACCACGCTGTCTACCTTCCGGAAGCTGTGTTGTGAGGCGATCTTGTTCTACAC 528
 Db 481 ACCACGCTGTCTACCTTCCGGAAGCTGTGTTGTGAGGCGATCTTGTTCTACAC 540
 QY 529 CAGGAAATCCGGAACATGTTTCCACCTGCGGCTTCTGTGACACAGACCTCCGACGCAAG 588
 Db 541 CAGGAAATCCGGAACATGTTTCCACCTGCGGCTTCTGTGACACAGACCTCCGACGCAAG 600
 QY 589 CTGTCTGAAAGATTTCTCCGGAAGCTGCGGACGCGCAGAGGAGGACCTGAGACAGATTCTGACG 648
 Db 601 CTGTCTGAAAGATTTCTCCGGAAGCTGCGGACGCGCAGAGGAGGACCTGAGACAGATTCTGACG 660
 QY 649 CAGTACACCACTTCTGTGAAGCCGGCTTTCAGAGATTCTGCTGCC----- 696
 Db 661 CAGTACACCACTTCTGTGAAGCCGGCTTTCAGAGATTCTGCTGCCCGCAGAGACATCTGA 720
 QY 697 -----ACMAAGATATG 709
 Db 721 CAGGGAATGAGTGCATGATGAGCCCATGATGTTGATGAGGGAACMAAGATATG 780
 QY 710 CCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGCAGC 769
 Db 781 CCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGCAGC 840
 QY 770 ACATCCAGACATTTGAAATGATGACATCTGCAATGACACCGAGAGGGTCAATGAGC 829
 Db 841 ACATCCAGACATTTGAAATGATGACATCTGCAATGACACCGAGAGGGTCAATGAGC 900
 QY 830 GGAAGTCAAGAGGAGACTTTTCTGAGCCAGGGGACCACTCTGGATGCTGACCTCTGGA 889
 Db 901 GGAAGTCAAGAGGAGACTTTTCTGAGCCAGGGGACCACTCTGGATGCTGACCTCTGGA 960
 QY 890 AACGTCACATTTGAGTCCAGACAGACCCCACTGAGAGGGCTGCGGACTCAAGGGA 949
 Db 961 AACGTCACATTTGAGTCCAGACAGACCCCACTGAGAGGGCTGCGGACTCAAGGGA 1020
 QY 950 GGTCTCCGCGCGCGCATGTGTGTTCAAGGACTGAGCTTGGGAGCGCCACCAACCCAC 1009
 Db 1021 GGTCTCCGCGCGCGCATGTGTGTTCAAGGACTGAGCTTGGGAGCGCCACCAACCCAC 1080

QY 1010 TGCTTCTCTCGGCGCACCCAGGGAGTGTAGCAGCGAGCCCTTCTCACTAGAGAGT 1069
 DB 1081 TGCTTCTCTCGGCGCACCCAGGGAGTGTAGCAGCGAGCCCTTCTCACTAGAGAGT 1140
 QY 1070 GGAACCTCAGATGTCTCACTAGACTCACTTGTCTGGACACTGACGAGCGTCTCTGAGG 1129
 DB 1141 GGAACCTCAGATGTCTCACTAGACTCACTTGTCTGGACACTGACGAGCGTCTCTGAGG 1200
 QY 1130 TTTTCAGCCACTTAGGCTGTTGCGGTTTAAAGATCCCTCAGGCACTGAGAAATGCA 1189
 DB 1201 TTTTCAGCCACTTAGGCTGTTGCGGTTTAAAGATCCCTCAGGCACTGAGAAATGCA 1260
 QY 1190 CAGAAATGTCAGGAAGCCTGGAGGCTTCTGTAGAAATGTGAGCACAATTATTTGGGAA 1249
 DB 1261 CAGAAATGTCAGGAAGCCTGGAGGCTTCTGTAGAAATGTGAGCACAATTATTTGGGAA 1320
 QY 1250 ATTGAGGAGACACCTTAACACTGCTGCTGCTGATGTTTGTGACATGAAACCAAGT 1309
 DB 1321 ATTGAGGAGACACCTTAACACTGCTGCTGCTGATGTTTGTGACATGAAACCAAGT 1380
 QY 1310 GGGAGAGATTTTTCAGTCTGATCTGTTCTTACACACTCACAATTAATCTAAAG 1369
 DB 1381 GGGAGAGATTTTTCAGTCTGATCTGTTCTTACACACTCACAATTAATCTAAAG 1436
 QY 1370 TTTTGTGAACAAGTACTTCTTTTATCATGTTACATGTCCTCATGTTTCTGTTTCT 1429
 DB 1437 TTTTGTGAACAAGTACTTCTTTTATCATGTTACATGTCCTCATGTTTCTGTTTCT 1496
 QY 1430 GTTTCATTAACACAGGCTGTTGCTTACAACTTAATTCATGACCCAGTGTGTTGC 1489
 DB 1497 GTTTCATTAACACAGGCTGTTGCTTACAACTTAATTCATGACCCAGTGTGTTGC 1556
 QY 1490 AGTCCAGGCTGAGCTTACAGGATATGGGAGCACTGAGGAGTGTTCCTCCCTGCTT 1549
 DB 1557 AGTCCAGGCTGAGCTTACAGGATATGGGAGCACTGAGGAGTGTTCCTCCCTGCTT 1616
 QY 1550 GTGCTTTAAGGCGACAGAGCGAGGCGATGCTCTGAAGACCCAGCATCACCAGG 1609
 DB 1617 GTGCTTTAAGGCGACAGAGCGAGGCGATGCTCTGAAGACCCAGCATCACCAGG 1676
 QY 1610 CTTGTGCGGGGCGAG 1624
 DB 1677 CTTGTGCGGGGCGAG 1691
 RESULT 5
 LOCUS BC015547 2072 bp mRNA linear PRI 04-OCT-2001
 DEFINITION Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9668
 IMAGE:3845821, mRNA, complete cds.
 ACCESSION BC015547
 VERSION BC015547.1 GI:15930229
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2072)
 TITLE Strausberg, R.
 JOURNAL Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-rt@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu

Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov
 Series: IRAP Plate: 20 Row: 5 Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.
 FEATURES
 source
 1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH_MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SportC"
 87..692
 /codon_start=1
 /product="Similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGEDCESPAPADRPHPPLIGVSGTASGKTVCEKIM ELIGONEVEBORRVVILSODREYKVLTEOKKALKGOYFNPDPAPNDLHRTLK NIVEGKVEVPTDVPVTHSLRPETTVVADVVLFGSLIVFSQELRDMFHLRLPVDT DSDVRLSRDKRRCRDHPFRSQYCHQDPRAAHGHSEW"
 BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN
 Query Match 86.0%; Score 1397.4; DB 9; Length 2072;
 Best Local Similarity 93.8%; Pred. No. 1,3e-284;
 Matches 1517; Conservative 0; Mismatches 1; Indels 99; Gaps 2;
 QY 8 CGCCTTCGACCTTGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGGA 67
 DB 1 CGCCTTCGACCTTGGCGCTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGGA 60
 QY 68 CCCGATGCGGGGAGCGGAGCGGAGATGCGCTTCCGCGGAGCGGAGACTGCGAGAGCC 127
 DB 61 CCCGATGCGGGGAGCGGAGCGGAGATGCGCTTCCGCGGAGCGGAGACTGCGAGAGCC 120
 QY 128 CCGGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 187
 DB 121 CCGGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 180
 QY 188 CTGCGACGGGAGAGTGCACCGTGTGTGAGAAATCATGAGATTCTGGGACAGAGCGAGG 247
 DB 181 CTGCGACGGGAGAGTGCACCGTGTGTGAGAAATCATGAGATTCTGGGACAGAGCGAGG 240
 QY 248 TGGAAACGGGCGGAGCGGAGGAGTGTGATCCTGAGCGAGGACAGGTTCTCAAGTCTCTGA 307
 DB 241 TGGAAACGGGCGGAGCGGAGGAGTGTGATCCTGAGCGAGGACAGGTTCTCAAGTCTCTGA 300
 QY 308 CGGACAGGCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 367
 DB 301 CGGACAGGCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 360
 QY 368 TTGATAATGATTTGATGACAGGACTTGAAGAATCTGTGAGGCGCAAAACGCTGAGAG 427
 DB 361 TTGATAATGATTTGATGACAGGACTTGAAGAATCTGTGAGGCGCAAAACGCTGAGAG 420
 QY 428 TGCAGACCTATGATTTTGTGACACACTCAAGGTTACCAAGAGACCAAGTGTGTTACCTTG 487
 DB 421 TGCAGACCTATGATTTTGTGACACACTCAAGGTTACCAAGAGACCAAGTGTGTTACCTTG 480
 QY 488 CGGACGCGGTTCTGTTTGAAGGATCTGTGTTCTTACAGCGAGGATCCGGGACATGT 547
 DB 481 CGGACGCGGTTCTGTTTGAAGGATCTGTGTTCTTACAGCGAGGATCCGGGACATGT 540
 QY 548 TCACCTGCGCTCTTCTGTGACACCGAGCTCCAGCTGTCTCGAAGAGTCTCC 607

```

Db 541 TCCACCTGCGCTCTCTGTCGACACCGATCCGACGTGACGCTGTCGAGA----- 593
Qy 608 GGGAGTGGCGCGGAGGAGGAGCCTGGAGCAGATTCTGACCGCAGTACACACCTTTCGTA 667
Db 594 ----- 593
Qy 668 AGCCGCGCTTCGAGGAGTTCTGCTGCCGACAAAGAAATATGCGGATGTGATCCAC 727
Db 594 -----GACAAAGAAATATGCCGATGTGATCCAC 625
Qy 728 GAGGATGACAAATATGTTGCCATCAACTGATGTGACAGCAATCCAGCAATTCTGA 787
Db 626 GAGGATGACAAATATGTTGCCATCAACTGATGTGACAGCAATCCAGCAATTCTGA 685
Qy 788 ATGGATCATCTGCAAAATGAGCAGGAGGAGGTCCAAATGGGCGGAGCTACAAAGGAGCCT 847
Db 686 ATGGATCATCTGCAAAATGAGCAGGAGGAGGTCCAAATGGGCGGAGCTACAAAGGAGCCT 745
Qy 848 TTTCTGAGCCAGGGGAGCACCCTGGGATGCTGACCTCTGGCAACGGTCAATTGGAGT 907
Db 746 TTTCTGAGCCAGGGGAGCACCCTGGGATGCTGACCTCTGGCAACGGTCAATTGGAGT 805
Qy 908 CCAAGACAGACACCCCACTGAGGGGCTGCCGAGCCTCAGGGGAGGATCTCCGCCGCGCATG 967
Db 806 CCAAGACAGACACCCCACTGAGGGGCTGCCGAGCCTCAGGGGAGGATCTCCGCCGCGCATG 865
Qy 968 TGTGTTAGGGAGCTGAGGCTGGGGAGCGCCCAACCCCACTGCTTCTCCGCGGCAC 1027
Db 866 TGTGTTAGGGAGCTGAGGCTGGGGAGCGCCCAACCCCACTGCTTCTCCGCGGCAC 925
Qy 1028 CCCAGGGGAGTGTAGAGCAGGAGCCTTCCCTCACTCAGGAGTGGAAACTCAGATGTGCA 1087
Db 926 CCCAGGGGAGTGTAGAGCAGGAGCCTTCCCTCACTCAGGAGTGGAAACTCAGATGTGCA 985
Qy 1088 CTCAGACTCAACTGCTGGGACACTGACAGGCGTTTCTGAGGTTTTCAGCCACTTAGGCT 1147
Db 986 CTCAGACTCAACTGCTGGGACACTGACAGGCGTTTCTGAGGTTTTCAGCCACTTAGGCT 1045
Qy 1148 CGTTGGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCACGAATGTGACGAAAGCC 1207
Db 1046 CGTTGGCGGTTTAAAGATCCCTCTAGGTCACTGAGAAATGCCACGAATGTGACGAAAGCC 1105
Qy 1208 TGGGAGGCTTCTGAGGAAATGTAGAGCACTAATTGGGGGAAATTGAGAGACAGCCTAG 1267
Db 1106 TGGGAGGCTTCTGAGGAAATGTAGAGCACTAATTGGGGGAAATTGAGAGACAGCCTAG 1165
Qy 1268 ACACTGCTGGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGAGTTCCTCA 1327
Db 1166 ACACTGCTGGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGAGTTCCTCA 1225
Qy 1328 GTCTGATCTGGTCTTACACTCAGCACTACACTAATCTGAAAGTTTGTGAAACAATGCTT 1387
Db 1226 GTCTGATCTGGTCTTACACTCAGCACTACACTAATCTGAAAGTTTGTGAAACAATGCTT 1281
Qy 1388 TCCCTTTTTCATGTTACATGCTCCTCATGTTTTCTGTTTTCTGTTTCAACAAGGCT 1447
Db 1282 TCCCTTTTTCATGTTACATGCTCCTCATGTTTTCTGTTTTCTGTTTCAACAAGGCT 1341
Qy 1448 GGTGTGGCTCAAAACCTAATTTTCAATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACA 1507
Db 1342 GGTGTGGCTCAAAACCTAATTTTCAATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACA 1401
Qy 1508 CGGATATGGGGAGCACTGAGGAGATGTTTTCCCGGCTTGTGCTTAAAGGCAAGA 1567
Db 1402 CGGATATGGGGAGCACTGAGGAGATGTTTTCCCGGCTTGTGCTTAAAGGCAAGA 1461
Qy 1568 AGGAGGCGGATGCTTGAAGACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 1624
Db 1462 AGGAGGCGGATGCTTGAAGACCCAGCATCAACCCAGGCTTGTGCGGGGCGAG 1518

```

RESULT 6
AF125106

```

LOCUS AF125106 1395 bp mRNA linear PRI 07-FEB-2002
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF125106
VERSION AF125106.1 GI:18568108
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Xie, Y.R., Yu, L., and Zhao, S.Y.
TITLE Cloning of a new human cDNA similar to Mus musculus uridine kinase
mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1395)
AUTHORS Ding, J.B., Yu, L., and Zhao, S.Y.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
People's Republic of China
FEATURES
source
location/Qualifiers
1..1395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
8..838
/codon_start=1
/product="uridine kinase"
/protein_id="P4L75943.1"
/db_xref="GI:18568109"
/translation="MASGSEGRARAGANRPHORPFLIGSGTASGKSTVEKIME
LLQONVEQHQKGVILISODRPFYKUTATGAKAKALQYNEFDPDAFNDLHRTLN
IVEGKVEVPTDVEVHSRLPETVYPAADVLFESILVFSQELDMFLRLVPTD
SDVRLSRVLRVLRGRDEQILTOYTTAPKAFEEFCLPTKKYAVIILPRGVNVA
INIVQIHDILINDICKWHRGSGNSRTYKTPSPBBDHPGMLTSGKSLHSSSRPH
"
BASE COUNT 325 a 352 c 411 g 307 t
ORIGIN
Query Match 71.2%; Score 1156.8; DB 9; Length 1395;
Best Local Similarity 94.1%; Pred. No. 8.6e-234;
Matches 1317; Conservative 0; Mismatches 57; Indels 26; Gaps 10;
Qy 87 GGGCGAGATGCTTGGCGGAGGCGGAGAGCTCCGAGAGCCCGCGGAGCCGACCG 146
Db 1 GGGCGAGATGCTTGGCGGAGGCGGAGAG---GTGGAAGGCCCGCGGAGCAGAACG 57
Qy 147 TCCGACCAAGCGGCTCTCTGATAGGGGTGAGCGGCGGACCTGCGAGGAGGAGCGAC 206
Db 58 TCCGACCAAGCGCTCTCTGATAGGGGTGAGCGGCGGACCTGCGAGGAGGAGCGAC 117
Qy 207 CGTGTGAGAAAGATCATGAGATGCTGGGACAGAAAGAGTGGAAACAGCGGACGCGAA 266
Db 118 CGTGTGAGAAAGATCATGAGATGCTGGGACAGAAAGAGTGGAAACAGCGGACGCTTA 177
Qy 267 GGTGTGATCTGAGGAGGAGGAGGAGGTTTCAAGAGTCTGACGCGGAGGAGGAGGCGCAA 326
Db 178 GGTGTGATCTGAGGAGGAGGAGGAGGTTTCAAGAGTCTGACGCGGAGGAGGAGGCGCAA 237
Qy 327 GGCCTTGAAGAGCAGTACAAATTTTGAACCATGAGTGCCTTTGATTAATGATTGATGCA 386
Db 238 GGCCTTGAAGAGCAGTACAAATTTTGAACCATGAGTGCCTTTGATTAATGATTGATGCA 297
Qy 387 CAGACTCTGAAGAACATGCTGAGAGGCAAAAAGTGAAGTGCAGCACTATGATTTGT 446
Db 298 CAGACTCTGAAGAACATGCTGAGAGGCAAAAAGTGAAGTGCAGCACTATGATTTGT 357
Qy 447 GACACCTAAGATTACAGAGACCAACGCTGTTACCTTGGGAGAGTGTCTGTTTGA 506
Db 358 GACACCTAAGATTACAGAGACCAACGCTGTTACCTTGGGAGAGTGTCTGTTTGA 417
Qy 507 GGGCATCTGTTGTTCTACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTTGGT 566

```

```

Db 418 GGGCATCTTGGTCTTCTACAGCCAGAGATCCGGGACATGTTCCACTCGCCCTTCTTGGT 477
Qy 567 GGACACCGCATCTCCGACGTCAGGCTGTCTCGAAGATTTTCCGGGACCTGCGCCCGAGGAG 626
Db 478 GGACACCGCATCTCCGACGTCAGGCTGTCTCGAAGATTTTCCGGGACGTCGCGCCCGAGGAG 537
Qy 627 GGACCTGGAGCAGATTTCTGACGACAGTACACCACTTCTGGAAGCCGCGCTTCGAGAGATT 686
Db 538 GGACCTGGAGCAGATTTCTGACGACAGTACACCACTTCTGGAAGCCGCGCTTCGAGAGATT 597
Qy 687 CTGCTGCTCCGACAAAGAGATATGCGGATGTATATCCACGAGAGATGACAAATATGAT 746
Db 598 CTGCTGCTCCGACAAAGAGATATGCGGATGTATATCCACGAGAGATGACAAATATGAT 657
Qy 747 TGCCATCAACCTGATCGTGACAGCATCATCGAGACATTCGTAATGGTGCATCTGAAATG 806
Db 658 TGCCATCAACCTGATCGTGACAGCATCATCGAGACATTCGTAATGGTGCATCTGAAATG 717
Qy 807 GCACCGAGAGAGGTCCTCAATGGCGGAGCTACAGCGGACCTTTTCTGACCGAGGAGCA 866
Db 718 GCACCGAGAGAGGTCCTCAATGGCGGAGCTACAGCGGACCTTTTCTGACCGAGGAGCA 777
Qy 867 CCTTGGATGCTGACCTCTGGAACCGGTCACTTTGAGTTCAGCAGACAGCCCACTG 926
Db 778 CCTTGGATGCTGACCTCTGGAACCGGTCACTTTGAGTTCAGCAGACAGCCCACTG 837
Qy 927 AGGGGCTGCGGACCTCAAGGGAGGTCCTCGCGCGGATGTGTAGGAGCTGAGCC 986
Db 838 AGGGGCTGCGGACCTCAAGGGAGGTCCTCGCGCGGATGTGTAGGAGCTGAGCC 897
Qy 987 TGGGACCGCCACCCACACCACTGCTTCTCTCGGCGACCCCGAGGAGTGTATGAG 1046
Db 898 TGGGACCGCCACCCACACCACTGCTTCTCTCGGCGACCCCGAGGAGTGTATGAG 957
Qy 1047 CGAGGCTTCTCTCACTCGAGAGTGAATCAAGATGTGTCACTCAACTCAACTTGTCTG 1106
Db 958 CGAGGCTTCTCTCACTCGAGAGTGAATCAAGATGTGTCACTCAACTCAACTTGTCTG 1016
Qy 1107 GACACTGACAGGCGCTTCTGAGGTTTTCAGGCACTTGAAGCTGTCGCGGTTTAAAGATCC 1166
Db 1017 GACACTGACAGGCGCTTCTGAGGTTTTCAGGCACTTGAAGCTGTCGCGGTTTAAAGATCC 1076
Qy 1167 CTCTAGGTAC--TGAGAAATGCAAG--AAATGTCAGAGAAAGCTG----GAGAGCTTC 1218
Db 1077 CTCTAGGTACCTTGAAGAAATGCAAGAAATGTCAGAGAAAGCTGCGGAAAGCTTCTG 1136
Qy 1219 TGTGAGAAATGAGGCACTTATTTGGGAAA-----TTGAGAGACAGCCTTAAGCACTG 1273
Db 1137 TGAAGGAATGTAAGGCACTTAAATTTGGGCAAAATTGAGAAAGACAGCCTTAAGCACTT 1196
Qy 1274 G---CTGGCTGATGTTTGTGACA--GTGAACCCAGAGTGG--AGAGAGTTTTCAG 1328
Db 1197 GGGCTTGGCTTAATTTTGTGACATGTGAACCCAGAGTGAAGAGATTTTTCAG 1256
Qy 1329 TCTGATCTGTTTCTTACACACTCAACAACATTAACAAAGTTTGTGAACAAGTACTT 1388
Db 1257 TCTGATCTGTTTCTTACA-----CACACACATTAACAAAGTTTGTGAACAAGTACTT 1312
Qy 1389 CCTTTTAAACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTG 1448
Db 1313 CCTTTTAAACATGTTACATGTCCTCATGTTTCTGTTTCTGTTTCAATACAAAGCTG 1372
Qy 1449 GTTGTGGCTTCAAAACCTAA 1468
Db 1373 GTTGTGGCTTCAATTA 1392

```

```

RESULT 7
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens uridine-cytidine kinase I (UCKI) mRNA, complete cds.
DEFINITION AF237290
ACCESSION AF237290.1 GI:13506764
VERSION

```

```

KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
JOURNAL human uridine-cytidine kinases
MEDLINE Mol. Pharmacol. 59 (5), 1181-1186 (2001)
21203813
PUBMED 11306702
REFERENCE 2 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
JOURNAL University Hospital, Stockholm 14186, Sweden
FEATURES
source
1. 1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 1022
/gene="UCKI"
5. 838
/gene="UCKI"
/function="phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate"
/codon_start=1
/product="uridine-cytidine kinase 1"
/protein_id="AAK28324.1"
/db_xref="GI:13506765"
/translacion="MASAGGDESPAPRAPHORPFLIGSGGTASGKSTCEKIM
ELIGONVEERQKRVITLSDRFYKTLTEQKAKLKGQYNDPRAFPDNDLHRLTK
NIVEGTVPEPYDFVTHSRLEPTTVVYADAVLFEGLIFVSQELRDMFHLFPDT
DSVRLSRVLRVRRGRDLEQLITQYTFVKAPEFELPTKXADVIIPREVDNV
AIIULYQHIQDILINDICIKWHRSGNSGRSKYTFSPDHPMLTGKSHLESSRP
H"
BASE COUNT 227 a 282 c 324 g 189 t
ORIGIN
Query Match 62.7%; Score 1018.8; DB 9; Length 1022;
Best Local Similarity 99.8%; Pred. No. 1.2e-204;
Matches 1020; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 90 CGAGATGCTTGGCGGGAGGCGAAGACCTGCGGAGCCCGGAGCGGACCTGCC 149
Db 1 CGAGATGCTTGGCGGGAGGCGAAGACCTGCGGAGCCCGGAGCGGACCTGCC 60
Qy 150 GCACAGCGGCGCTTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGAGTCAGCGT 209
Db 61 GCACAGCGGCGCTTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGAGTCAGCGT 120
Qy 210 GTGTGAAGATCATGAGTGTCTGGAACAAGAGGTGAACAGCGGCGAGGAGGT 269
Db 121 GTGTGAAGATCATGAGTGTCTGGAACAAGAGGTGAACAGCGGCGAGGAGGT 180
Qy 270 GTTCATCTGAGCGCAGAGACAGGTTCTCAAGAGTCTGACGGCAGACAGGCGCAAGGC 329
Db 181 GTTCATCTGAGCGCAGAGACAGGTTCTCAAGAGTCTGACGGCAGACAGGCGCAAGGC 240
Qy 330 CTTGAAGAGCAGTACAAATTTGACATCCAGATGCTTTGATATGATTTGACAG 389
Db 241 CTTGAAGAGCAGTACAAATTTTGAATCCAGATGCTTTGATATGATTTGACAG 300
Qy 390 GACTGGAAGACATGATGAGGCGAAAACGATGAGGTGCGGACCTATGATTTGTGAC 449

```


[illegible]

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Ch9>

RP11-33406 is from the library RPEC-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

This sequence is the entire insert of clone RP11-33406. The true left end of clone RP11-40A7 is at 113870 in this sequence. The true right end of clone RP11-643B14 is at 63282 in this sequence.

FEATURES		Location/Qualifiers	
source	1..147492		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="9"		
	/clone="RP11-334J6"		
	/clone_1ib="RP11-11.2"		
	30143..30199		
	/note="Sequence from overlapping clone RP11-643E14 (AL554855). Assembly confirmed by restriction digest."		
BASE COUNT	32992 a 36290 c 37577 g 40633 t		
ORIGIN			
Query Match	54.2% Score 880; DB 9; Length 147492;		
Best Local Similarity	100.0%; Pred. No. 2.4e-175;		
Matches	880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	745 GTTGCATCAACCTGATCGTGCAGACATCCAGACATTCTGAATGTGATCATCTGCAAA	804	
Db	129488 GTTGCATCAACCTGATCGTGCAGACATCCAGACATTCTGAATGTGATCATCTGCAAA	1294289	
Qy	805 TGGCACCCAGAGAGGCTCCAAATGGGGGAGCTCAAGCCGACCTTTTCTAGCCAGGGGAC	864	
Db	129428 TGGCACCCAGAGAGGCTCCAAATGGGGGAGCTCAAGCCGACCTTTTCTAGCCAGGGGAC	1293659	
Qy	865 CACCCCTGGATGTCTGACCTCTTGCCAAAGGTGCACATTTGGAATCCAGCAGACAGACCCAC	924	
Db	129368 CACCCCTGGATGTCTGACCTCTTGCCAAAGGTGCACATTTGGAATCCAGCAGACAGACCCAC	129309	
Qy	925 TGAGGGGCTGCCAGACTCTCAGGGCAGGTCCTCCCGCCGATGTGTTCAGGGACTGAG	984	
Db	129308 TGAGGGGCTGCCAGACTCTCAGGGCAGGTCCTCCCGCCGATGTGTTCAGGGACTGAG	129249	
Qy	985 CCTGGGAGCGCCACCCACCACTCGCTCTCCCTCGGGGACCCCAAGGGGAGTGTAGC	1044	
Db	129248 CCTGGGAGCGCCACCCACCACTCGCTCTCCCTCGGGGACCCCAAGGGGAGTGTAGC	1291888	
Qy	1045 AGCGAGGCTTCTCTCACTCAGAGTGGAACTCAGATGTGTCACTCAGACTCAACTTGGCT	1104	
Db	129188 AGCGAGGCTTCTCTCACTCAGAGTGGAACTCAGATGTGTCACTCAGACTCAACTTGGCT	129129	
Qy	1105 GGGACACTGACAGAGGTTCTCGAGGTTTTCAGGCCTTAAGGTCGTCGGGTTTAAAGAT	1164	
Db	129128 GGGACACTGACAGAGGTTCTCGAGGTTTTCAGGCCTTAAGGTCGTCGGGTTTAAAGAT	1290659	
Qy	1165 CCTCTAGTCACTGAGAAATGCCACAGAAATGTGAGGAAGCTTGAGAGCTTCTGTGAG	1224	

Db 129068 CCTCTAGCTACGTGAGAAATGCCACAGAACTGTGAGAGAACCTGGAGGCTTCTGTGAG 129009
Qy 1225 GAATGTGAGGCACATTTATTGGGAAATTTGAGAGACAGCCTAGACACTGGCTGGCTGAT 1284
Db 129008 GAAATGTGAGGCACATTTATTGGGAAATTTGAGAGACAGCCTAGACACTGGCTGGCTGAT 128949
Qy 1285 GTTTTGTGAGAGGAAACCCACAGTGGAGAGAGTTTTCAGAGCTGATCTGGTCTTA 1344
Db 128948 GTTTTGTGAGAGGAAACCCACAGTGGAGAGAGTTTTCAGAGCTGATCTGGTCTTA 128889
Qy 1345 CACACTCAGACATTAACCTCAAAAGTTTGTGAACAGTAATCTTTTAAATGTTA 1404
Db 128888 CACACTCAGACATTAACCTCAAAAGTTTGTGAACAGTAATCTTTTAAATGTTA 128829
Qy 1405 CATGTCCATGATTTTCTGTTTCTGTTTCAATACAGAGCTGGTTGGCTTAACAAC 1464
Db 128828 CATGTCCATGATTTTCTGTTTCTGTTTCAATACAGAGCTGGTTGGCTTAACAAC 128769
Qy 1465 CTAATTTTCATGACCCAGTGTGTTGCAATCCAGCGCTCAACAGATATGGGAGCCAC 1524
Db 128768 CTAATTTTCATGACCCAGTGTGTTGCAATCCAGCGCTCAACAGATATGGGAGCCAC 128709
Qy 1525 TGAGGATGTTTTTCCCTTCTGTTGCTTAAAGCAGAGAAAGCGAGATGCCCT 1584
Db 128708 TGAGGATGTTTTTCCCTTCTGTTGCTTAAAGCAGAGAAAGCGAGATGCCCT 128649
Qy 1585 GGAAGCACCCAGCATCAACCCAGGCTGTGCGGGGCCAG 1624
Db 128648 GGAAGCACCCAGCATCAACCCAGGCTGTGCGGGGCCAG 128609

RESULT 9
AX449219 834 bp DNA linear PAT 03-JUL-2002
LOCUS AX449219
DEFINITION Sequence 3 from Patent WO0202761.
ACCESSION AX449219
VERSION AX449219.1 GI:21697996
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Glucksmann, M. A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 3 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1..834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN

Query Match 51.4%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 1.3e-165;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 ATGGCTTGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 153
Db 1 ATGGCTTGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
Qy 154 CAGCGGCCCTTCTGATGAGGGGTGAGCGGCGCACTGCGACCGGAAAGTGCAGCTGTGT 213
Db 61 CAGCGGCCCTTCTGATGAGGGGTGAGCGGCGCACTGCGACCGGAAAGTGCAGCTGTGT 120
Qy 214 GAGAAGATCATGAGATTCTGGGACAGAAAGAGGTGAAACAGCGGCGAGCGGAAGTGGTC 273
Db 121 GAGAAGATCATGAGATTCTGGGACAGAAAGAGGTGAAACAGCGGCGAGCGGAAGTGGTC 180
Qy 274 ATCTGAGCCAGGAGAGGTTCTCAAGAGTCTTGAACGCGAGAGCAGAAAGGCCCTTG 333

Db 181 ATCTGAGCCAGGACAGGTTCTAAGAGTCTCTACCGGACAGAGCAGAAAGGCCCTTG 240
Qy 334 AAAGACAGTGAATTTTGAACATCCAGATGCTTTGATATATGATTTGATGACAGGACT 393
Db 241 AAAGACAGTGAATTTTGAACATCCAGATGCTTTGATATATGATTTGATGACAGGACT 300
Qy 394 CTGAGAAACATCGTGAAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGAACAC 453
Db 301 CTGAGAAACATCGTGAAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGAACAC 360
Qy 454 TCAAGTTACAGAGACCAAGTGTGTCTACCTGCGGACGTGTTCTTTGAGGGCATC 513
Db 361 TCAAGTTACAGAGACCAAGTGTGTCTACCTGCGGACGTGTTCTTTGAGGGCATC 420
Qy 514 TTGGTGTCTACAGCCAGGAGATCCGGAGATGTTCACTGGCGCTCTTCTGAGAC 573
Db 421 TTGGTGTCTACAGCCAGGAGATCCGGAGATGTTCACTGGCGCTCTTCTGAGAC 480
Qy 574 GACTCGAGCTCAGGCGTGTCTGAGAGTCTCCGGAGAGTGGCGGAGGAGGACCTG 633
Db 481 GACTCGAGCTCAGGCGTGTCTGAGAGTCTCCGGAGAGTGGCGGAGGAGGACCTG 540
Qy 634 GAGCAGATTCTGACGCGATACACACCTTCTGTAAGCGGCTTTCAGAGAGTTCTGCTG 693
Db 541 GAGCAGATTCTGACGCGATACACACCTTCTGTAAGCGGCTTTCAGAGAGTTCTGCTG 600
Qy 694 CCGACAAAGAAATGATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 753
Db 601 CCGACAAAGAAATGATGCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
Qy 754 AACCTGATGTGAGCAGCATCCAGGACATTCCTGAATGTGACATCTGCAATGGACCGA 813
Db 661 AACCTGATGTGAGCAGCATCCAGGACATTCCTGAATGTGACATCTGCAATGGACCGA 720
Qy 814 GAGAGGTCCAAATGGCGGAGCTCAAAACGGGACCTTTCTGAGCCAGGAGACACCTGGG 873
Db 721 GAGAGGTCCAAATGGCGGAGCTCAAAACGGGACCTTTCTGAGCCAGGAGACACCTGGG 780
Qy 874 ATGCTGACCTTGTGGCAAAACGCTCAATTTGGATCCAGAGCAGACCCCACTGA 927
Db 781 ATGCTGACCTTGTGGCAAAACGCTCAATTTGGATCCAGAGCAGACCCCACTGA 834

RESULT 10
AF254133 834 bp mRNA linear PRI 02-MAY-2001
LOCUS AF254133
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Ho, Y.S. and Johnson, R.K.
TITLE Human uridine kinase from prostate cancer cell line (LNCap)
JOURNAL Unpublished
TITLE Direct Submission
JOURNAL Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA
FEATURES
source 1..834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"


```

/protein_id="AAK49122.1"
/db_xref="GI:13924750"
/translation="MASAGEPDESPAPREADRPHORPFLIGSGSTRASGKTYCEKIM
ELIGQNEVEORRQKVVILSDRFKVLTAQKALMGQNFHPDFFDNDLHRTLK
NIVEGKTEVEPTDVFTHSLPETTVYPPADVLPEGLVYSQELRDMHLLFVDT
DSDEVRLSRVLRGRDLQILTYQTVTKAEFLPTKQADAVIIPRGDMNV
AINDIVQHIQDILNDGICKWHRRGSGNSRKYKTFSEBGRHPMLTSKSRHLSSSRP
H"
BASE COUNT      194 a      219 c      265 g      156 t
ORIGIN
Query Match      51.3%; Score 832.4; DB 9; Length 834;
Best Local Similarity 99.9%; Pred. No. 2.9e-165;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
94 ATGGCTTGGCGGGAGGCGAAGACTGCGAAGACCCCGCGCGAGCGGACCGTCCGAC 153
1 ATGGCTTGGCGGGAGGCGAAGACTGCGAAGACCCCGCGCGAGCGGACCGTCCGAC 60
154 CAGCGGCGCTTCTCGATAGGGGGTGAAGGGGCGCATCGCCAGCGGGAAGTGCACCTGTGT 213
61 CAGCGGCGCTTCTCGATAGGGGGTGAAGGGGCGCATCGCCAGCGGGAAGTGCACCTGTGT 120
214 GAGAAATCATGAGATTCTGCGACAGAACGAGGTGGAACAGCGCGACCGGAAGTGTCT 273
121 GAGAAATCATGAGATTCTGCGACAGAACGAGGTGGAACAGCGCGACCGGAAGTGTCT 180
274 ATCTTGAGCCAGACAGAGTCTTCAAGAGTCTTGACGCGACAGACGAAGCCGCTTG 333
181 ATCTTGAGCCAGACAGAGTCTTCAAGAGTCTTGACGCGACAGACGAAGCCGCTTG 240
334 AAGGAGATGACAAATTTGACCATCCAGATCCCTTATATATGATTTGAGGACAGACT 393
241 AAGGAGATGACAAATTTGACCATCCAGATCCCTTATATATGATTTGAGGACAGACT 300
394 CTGAAAGAACATCGTGAAGGCGAAACCGTGGAGGTGCGACCTATGATTTTGTGACAC 453
301 CTGAAAGAACATCGTGAAGGCGAAACCGTGGAGGTGCGACCTATGATTTTGTGACAC 360
454 TCAAGGTTACAGAGACACAGGTGTCTACCTGCGGACGTTGTTGAGGGATC 513
361 TCAAGGTTACAGAGACACAGGTGTCTACCTGCGGACGTTGTTGAGGGATC 420
514 TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCACTGCGCTCTTCTGTGGAAC 573
421 TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCACTGCGCTCTTCTGTGGAAC 480
574 GACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGGAAGGACCTG 633
481 GACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGGAAGGACCTG 540
634 GAGCAGATTCTGACGACGATACACCACTTGTGAAGCGGCGCTTGAAGAGTTCTGCTG 693
541 GAGCAGATTCTGACGACGATACACCACTTGTGAAGCGGCGCTTGAAGAGTTCTGCTG 600
694 CCGACAAAGAGATGCGCATGTGATATCCCAAGAGAGTGAACAATATGTTGCGATC 753
601 CCGACAAAGAGATGCGCATGTGATATCCCAAGAGAGTGAACAATATGTTGCGATC 660
754 AACCTGATGTCGACATTCAGAGACATTCGATGATGTCATCTGCAATGGACCGA 813
661 AACCTGATGTCGACATTCAGAGACATTCGATGATGTCATCTGCAATGGACCGA 720
814 GAGAGGTCATGAGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGGG 873
721 GAGAGGTCATGAGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGGG 780
874 ATGCTGACCTCTGGCAACGCTCAATTTGAGAGTCGACGACGACGACCCCACTGA 927
781 ATGCTGACCTCTGGCAACGCTCAATTTGAGAGTCGACGACGACGACCCCACTGA 834

```

```

BD146824
LOCUS      BD146824      753 bp      DNA      linear      PAT 17-JUN-2003
DEFINITION
Primer for synthesizing full-length cDNA and use thereof.
ACCESSION
BD146824
VERSION
BD146824.1 GI:27852582
KEYWORDS
JP 2002191363-A/1667.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
REFERENCE
Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Iehi,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1667 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS      Homo sapiens (human)
PN      JP 2002191363-A/1667
PD      09-JUL-2002
PF      28-JUL-2000 JP 2000280990
PI      TOSHIO OTA, TAKAO ISOgai, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI      SATTO,
PI      JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI      KEIICHI NAGAI, TETSUJI OTSUKI
PC      C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC      10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC      Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
FT      source      1..753
FT      Location/Qualifiers
FT      source      1..753
FT      Location/Qualifiers
FT      source      1..753
FT      Location/Qualifiers
BASE COUNT      158 a      196 c      261 g      135 t      3 others
ORIGIN
Query Match      43.7%; Score 710; DB 6; Length 753;
Best Local Similarity 97.9%; Pred. No. 2.1e-139;
Matches 729; Conservative 0; Mismatches 13; Indels 3; Gaps 1;
6 GTGCTCCGACCTCGGCGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 65
7 GTGCTCCGACCTCGGCGTGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 66
66 GACCCGATGCGGGGAGCGGAGCGGAGTGGCTTGCGGGGAGCGGAGACGCGAGAG 125
67 GACCCGATGCGGGGAGCGGAGCGGAGTGGCTTGCGGGGAGCGGAGACGCGAGAG 126
126 CCCCGCGGAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 185
127 CCCCGCGGAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 186
186 CACTGCCAGGGGAGTGCACCGTGTGGAAGATCATGAGTCTCTGGACAGACGA 245
187 CACTGCCAGGGGAGTGCACCGTGTGGAAGATCATGAGTCTCTGGACAGACGA 246
246 GGTGGAACGCGGACCGGAGAGGTGTCATCTGAGCCAGGACAGGTTCTACAGGTCCT 305
247 GGTGGAACGCGGACCGGAGAGGTGTCATCTGAGCCAGGACAGGTTCTACAGGTCCT 306
306 GACGCGAGACAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 365
307 GACGCGAGACAGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 366
366 CTTTGAATATGATTTGATGACAGACTCTGAAGAACATCTGAGAGGCGAAGACGCTGA 425
367 CTTTGAATATGATTTGATGACAGACTCTGAAGAACATCTGAGAGGCGAAGACGCTGA 426
426 GGTGCGGACCTATGATTTTGTGACACATCAAGGTTACAGAGACCAAGGTGTCTACCC 485

```

```

Db      427  GGTGCGACATATGATTTTGTGACACACTCAAGTTACCAAGACCAACGGTGTCTACCC 486
Oy      486  TGGCGACGTGCTTCTGTTTGAAGGCACTTGTGTTCTTACACGCCAGGAGATCCGGACAT 545
Db      487  TGGCGACGTGCTTCTGTTTGAAGGCACTTGTGTTCTTACACGCCAGGAGATCCGGACAT 546
Oy      546  GTTCCACGTGCGCCCTCTCTGTTGACACCACTCCGACGTCAAGGTGTCTCGAAGATCT 605
Db      547  GTTCCACGTGCGCCCTCTCTGTTGACACCACTCCGACGTCAAGGTGTCTCGAAGATCT 606
Oy      606  CCGGACGTGCGCCCGAGGAGGAGCACTTGAAGCAATTTGACGAGTACACACACTTGT 665
Db      607  CCGGACGTGCGCCCGAGGAGGAGCACTTGAAGCAATTTGACGAGTACACACACTTGT 666
Oy      666  GAAGCGGCGCTTCCAGAGAGTCTGCGC---TGCAGCAAGAGATGCCATGTGATCAT 722
Db      667  GAAGCGGCGCTTCCAGAGAGTCTGCGC---TGCAGCAAGAGATGCCATGTGATCAT 726
Oy      723  CCCACGAGAGTGCACATATGTT 747
Db      727  TCCACGAGAGTGCACATATGANT 751

RESULT 12
LOCUS   BC025146 1959 bp mRNA linear ROD 16-APR-2003
DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cdna clone
ACCESSION BC025146
VERSION   BC025146.1 GI:19263563
KEYWORDS   MGC.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1959)
AUTHORS   Klausner R.D., Collins F.S., Wagner L.H., Derge J.G.,
           Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
           Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
           Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
           Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
           Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
           Scheetz T.E., Brownstein M.J., Uedin T.B., Tishiyuki S.,
           Abramson P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
           Abramson R.D., Mullany S.J., Bosak S.A., McEwen P.J.,
           McEwen K.J., Malek U.A., Gunaratne P.H., Richards S.,
           Wille K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,
           Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
           Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S.,
           Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
           Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
           Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
           Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalys D.E.,
           Scherch A., Schein J.E., Jones S.J. and Marra M.A.
           Generation and initial analysis of more than 15,000 full-length
           human and mouse cDNA sequences
           Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  22388257
MEDLINE  12477932
PUBMED   2 (bases 1 to 1959)
REFERENCE Strausberg R.
           Direct Submission
           Submitted (05-MAR-2002) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Jeffrey E. Green, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

```

DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 61 Row: p Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 1959
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_id="NCI CGAP_SG2"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
1. 1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translacion="MASAGGSGSEAPADRPORPPFLIGVSGTASGKSTCEKIM
ELIGQNEVPROROKVILSDCPYKLTLEOKAKLKGQVNDHPAPNDPMLHKLTK
NIEGTGTVPTPDPTVTRSLRLEPTVYADVLPEGLVFTVQERLDMFHLRPLDT
DSVIVSRKVLKRVQGRDEQLTYTFAVPAFEELPRTKADVLIPKVDNMV
AHLNLYQHLQDILNGDLCKRHRRGPNRHNKTFPPGDPGLATGKSHLESSRP
H"

BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Query Match 41.0%; Score 666.6; DB 10; Length 1959;
Best Local Similarity 84.4%; Pred. No. 3e-130;
Matches 750; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

Oy      47  GGGAGGGGGGGCGGGGGGACCCGATGCGCGGAGACGGCGCGAGTGGCTTGGCGG 106
Db      10  GGGCGGAGCGGCGCAAGTGTGGTAATGTATGTCTGCTCGGATATGAGCTTGGCGG 69
Oy      107  GAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGCGACGTCGCGACACGAGCGCCCTTCC 166
Db      70  GAGGCGGCGGCTCGAGAGAGCGCGCGCGCGAGGCGCATGTCTCCGACCGCGCGCTTCC 129
Oy      167  TGATAGGGGTGAGCGCGCGGCACTGCGAGCGGGAATGCAACCGTGTGGAAGATCATGG 226
Db      130  TCATCGCGGTGAGCGCGCGGCGGACCGCTAATGTGGCAACAGTGTGTGAGAAATCATGG 189
Oy      227  AGTTGCTGGGAGCAAGACGAGTGGAAACAGCGGAGGGAAGGTGCTCATCTGAGCGAG 286
Db      190  AGCTGCTGGGAGCAAGACGAGTGGACCGCGCGGAGCGCAAGTTGTCTATCTGAGCGAG 249
Oy      287  ACAGTTTCTTCAAGAGTCTGACGCGAGACAGAGCGCAAGGCTTTGAAGAGCACTACA 346
Db      250  ACAGTCTTCAAGAGTCTGACGCGCGAGACAGAGCGCAAGGCTTTGAAGAGCACTACA 309
Oy      347  ATTTTACCATCCAGATGCGCTTTGATTAATGATTGTGACACAGAGACTTGAAGACATCG 406
Db      310  ATTTTACCAACCCAGATGCTTTTGAATAATGATCTGTATGACAAAGACCTGAAAAAATTG 369

```

QY 407 TGGAGGCAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACTCAAGGTTACAG 466
DB 370 TTGAAGGCAAAACGCTGAGAGTGCCGACCTATGATTTTGTGACACTCAAGGTTACAG 429
QY 467 AGACCAAGTGTCTACCTGCGGACGTGTTCTGTTTGAAGGATCTTGTTTACCA 526
DB 430 AGACCAAGTGTCTACCTGCGGACGTGTTCTGTTTGAAGGATCTTGTTTACCA 489
QY 527 GCCAGAGATCCGGGAGATGTTCCACTGCGGCTCTTCGTGAGACACGACCTCCGAGCTCA 586
DB 490 CCCAGAGATCCGGGAGATGTTCCACTGCGGCTCTTCGTGAGACACGACCTCCGAGCTCA 549
QY 587 GCGTGTCTGAAAGATCTCCGGGACGTGCGCCGAGGAGGAGACCTGAGACATTTCTGA 646
DB 550 GCGTGTCTGAAAGATCTCCGGGACGTGCGCCGAGGAGGAGACCTGAGACATTTCTGA 609
QY 647 CGGAGTACACACCTTCTGTGAAGCCGCTTCGAGAGATTCGTGCTTCCGACAAAGAGT 706
DB 610 CTCAGTACACACCTTCTGTGAAGCCGCTTCGAGAGATTCGTGCTTCCGACAAAGAGT 669
QY 707 ATGCCCATGTGATCATCTCCACGAGAGTGAACAATATGTTGCCATCACTGATCGTGC 766
DB 670 ACCCTGATGATCATCTCCGAGAGTGAATATATGTTGCCATCACTGATCGTGC 729
QY 767 AGCAGATCCAGACATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 826
DB 730 AACACATCCAGACATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 789
QY 827 GCGGAGCTTCAAGCGGACCTTTTCTGAGCAGAGGAGCAACCTCGGAGTCTGACCTCTG 886
DB 790 GCGGAGCTTCAAGCGGACCTTTTCTGAGCAGAGGAGCAACCTCGGAGTCTGACCTCTG 849
QY 887 GCAAGGCTTCAATTTGAGTCCAGCAGCAGACCCCACTGAGGGGCTGC 935
DB 850 GCAAGGCTTCAATTTGAGTCCAGCAGCAGACCCCACTGAGGGGCTGC 898

RESULT 13
MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
LOCUS MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
DEFINITION Mus musculus uridine kinase mRNA, partial cds.
ACCESSION U11783
VERSION U11783.1 GI:471980
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1810)
AUTHORS Traut, T.W.
TITLE The functions and consensus motifs of nine types of peptide
segments that form different types of nucleotide-binding sites
JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE 94259063
PUBMED 8200357
REFERENCE 2 (bases 1 to 1810)
AUTHORS Kopp, P.A. and Traut, T.W.
TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse
JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE 97108719
PUBMED 8951040
FEATURES
source location/Qualifiers
1..1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
/BC_number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"

CDS
1..783
/BC_number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"

/protein_id="AAB50568.1"
/db_xref="GI:471981"
/translation="RQPRPFLIGSGTASGKSYCEKIMELLEGONEVDRORKLVI
LSQDCRYKVLAEQAKALKGQINPDPAPFNDDLMHKLKNIIVEKIVEPYPDVT
HSRLPEYTVVYPAADVLFEGILVFTYQELRDMFHLFLFDTSDFVLSRRLADVQK
RDLEQILQYTAFAVFAFEFCLPTKRYADVIIPRVDMVVAIINLVQHIQDILNBDL
CKRHGGPBNHKKRTPFERGDHPGLAAGKSHLESSSRPH"
misc_binding
join(37..63,412..426,451..465)
/bound_molecy="ATP at catalytic site"
polya_signal
/bound_molecy="ATP at catalytic site"
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN
Query Match 38.4%; Score 623; DB 10; Length 1810;
Best Local Similarity 86.7%; Pred. No. 4,8e-121;
Matches 686; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 145 CGTCCGACACGCGGCTTCTGATAGAGGAGCGGCGGCACTCCAGCGGAGAGTCG 204
DB 1 CGTCCGACACGCGGCTTCTGATAGAGGAGCGGCGGCACTCCAGCGGAGAGTCG 60
QY 205 ACCGTGTGAGAAATCATGAGTTCTGGGACAGAACGAGGTGGAACAGCGGCAAGCCG 264
DB 61 ACAGTGTGAGAAATCATGAGTTCTGGGACAGAACGAGGTGGAACAGCGGCAAGCCG 120
QY 265 AAGGTGTGATCCTGAGCGGACGAGTCTCAAGGTCTGACCGGACAGAGAGGCG 324
DB 121 AAGTGTGATCCTGAGCGGACGAGTCTCAAGGTCTGACCGGACAGAGAGGCG 180
QY 325 AAGGCTTGAAGAGACATTAATTTTGAACATCCAGATGCTTGTATATGATTTGATG 384
DB 181 AAGGCTTGAAGAGACATTAATTTTGAACATCCAGATGCTTGTATATGATTTGATG 240
QY 385 CACAGGACTCTGAAGAAATCTGTGAGGCGAAACGCTGAGGTGCCACTATGATTTT 444
DB 241 CACAGGACTCTGAAGAAATCTGTGAGGCGAAACGCTGAGGTGCCACTATGATTTT 300
QY 445 GTGACACTCAAGTTACAGAGACAGACAGTGTGTTACCTCGGAGCTGTTCTGTT 504
DB 301 GTGACACTCAAGTTACAGAGACAGACAGTGTGTTACCTCGGAGCTGTTCTGTT 360
QY 505 GAGGCGATCTGAGTGTCTACAGCCAGAGATCCGAGACATGTTCCAGCTGCGCTCTTC 564
DB 361 GAGGCGATCTGAGTGTCTACAGCCAGAGATCCGAGACATGTTCCAGCTGCGCTCTTC 420
QY 565 GTGACACGACTCCGACGCTGAGCTGTCTGAGAGTCTCCGAGACGTCGCGGAGGG 624
DB 421 GTGACACGACTCCGACGCTGAGCTGTCTGAGAGTCTCCGAGACGTCGCGGAGGG 480
QY 625 AGGAGCTTGAGAGCATTTCTGACGCACTACCACTTCTGTGAAGCGGCTTCCAGAGG 684
DB 481 AGGAGCTTGAGAGCATTTCTGACGCACTACCACTTCTGTGAAGCGGCTTCCAGAGG 540
QY 685 TTCTGCTCCGCAAAAGAGTATGCGATGATCATCCACGAGAGTGAACATATG 744
DB 541 TTCTGCTCCGCAAAAGAGTATGCGATGATCATCCACGAGAGTGAACATATG 600
QY 745 GTTGCCATCAACCTGATCTGAGCAGCATCCAGACATTTCTGAATGTGACATCTTCCAAA 804
DB 601 GTTGCCATCAACCTGATCTGAGCAGCATCCAGACATTTCTGAATGTGACATCTTCCAAA 660
QY 805 TGGACCCGAGAGAGGTTCAAATGGGCGGAGCTTCAAGCGGACCTTTTCTAGCCAGAGGAC 864
DB 661 TGGACCCGAGAGAGGTTCAAATGGGCGGAGCTTCAAGCGGACCTTTTCTAGCCAGAGGAT 720
QY 865 CACCTGTGATGATCTCTGTGCAAAACGCTGACATTTGAGTCCAGCAGAGACCCGAC 924
DB 721 CACCTGTGATGATCTCTGTGCAAAACGCTGACATTTGAGTCCAGCAGAGACCCGAC 780
QY 925 TGAAGGGCTGC 935
DB 781 TGAAGGGCTGC 791

```

RESULT 14
MUSURKIA          952 bp   DNA      linear      ROD 05-MAR-2001
LOCUS             Mus musculus uridine kinase pseudogene.
ACCESSION         L31784
VERSION           L31784.1 GI:471982
KEYWORDS          pseudogene; uridine kinase.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mus musculus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS           Traut, T.W.
TITLE             The functions and consensus motifs of nine types of peptide
                  segments that form different types of nucleotide-binding sites
                  1 (bases 1 to 952)
JOURNML           Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE           94259063
PUBMED            8200357
REFERENCE
AUTHORS           Ropp, P.A. and Traut, T.W.
TITLE             Cloning and expression of a cDNA encoding uridine kinase from mouse
                  brain
JOURNML           Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE           97108719
PUBMED            8951040
FEATURES
source            1..952
                  Location/Qualifiers
                  /organism="Mus musculus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10090"
                  /issue_type="kidney"
                  /ref. 913
                  /standard_name="uridine-cytidine kinase"
                  /pseudo
                  /codon_start=1
                  /product="uridine kinase"
                  misc_difference 130
                  /note="insertion not found in L31783"
BASE COUNT       230 a      231 c      277 g      214 t
ORIGIN
Query Match      36.8%; Score 597.4; DB 10; Length 952;
Best Local Similarity 82.3%; Pred. No. 1.2e-115;
Matches 699; Conservative 0; Mismatches 146; Indels 4; Gaps 1;
QY               91 GAGATGGCTTGGGGGGAGCGGAGACTGCGAGAGCCCGC-----GCCGAGGCCGACCG 146
Db               73 GAGATGGCTTGGGGGGAGCGGAGACTGCGAGAGTACTGCGCTGCCCGAGCGCAATCT 132
QY               147 TCCGACACGAGCGCCCTTCTCTGATAGGGGTGAGCGGGGCACTCCAGCGGAGAGTGCAC 206
Db               133 TCCGACAGCCTTGGCGCTTCTCTCATCGGGGTGAGTGGGGGCGCGCTGTGTGCAAGTCAAC 192
QY               207 CGTGTGAGAAATCATGAGTTGCTGGGACAGAGAGGTGGAACAAGCGGAGCGGAA 266
Db               193 AGTGTGAGAAATCATGAGTCTGGGACAGAAATGAAGTGAGACTCCGGTAGCCCAA 252
QY               267 GGTGGTATCCTGAGCGAGAGCAGGTTCTCAAGGTCCTGACGGGACAGAGAGAGGCGAA 326
Db               253 GTTGGTATCCTGAGCGAGAGCAGTCTTCTTCAAGGTTCTGACGGGACAGAGAGGCGAA 312
QY               327 GGCCTTGAAGAAGACATATTTTGAACATCCAGATGCCCTTGTATGATTTGATGCA 386
Db               313 GGCCTTGAAGAAGACATATTTTGAACATCCAGATGCCCTTGTATGATTTGATGCA 372
QY               387 CAGGACTCTGAAGAAATCGTGAAGGGGAAAACGGTGAAGGTGCGACCTATGATTTGT 446
Db               373 CAGGACTCTGAAGAAATCATTTATGAAGGCAAAAGTTGTAAGGTTCTTACATATGATTTGT 432
QY               447 GACACACTCAAGGTTTACAGAGACAGCGGTGCTTACCCTGCGAGAGTGTCTGTTGA 506
Db               433 GACGACTCAAGGTTTACAGAGACAGCTGTGTCTTACACATATGATGCGGTTCTGTTGA 492

```

```

QY               507 GGGCATCTTGATGTTTCTACAGAGATCCGGAGACATGTCACCTGCGCCTTCTGT 566
Db               493 GGGCATCTTGATGTTTCTACAGAGATCCGGAGACATGTTTCTACCTGCGCCTTCTGT 552
QY               567 GAGACACGACTCCGACGTGAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGGAGGAG 626
Db               553 GAGACACGACTCCGATGATTTAGGCTGTCTTGAAGAGTTCTCCGGGAGTGTGCAAGAGAG 612
QY               627 GAGACTGGACAGATTTCTGACGAGTACACACACTTCTGTAAACCGGCTTCGAGAGTT 686
Db               613 GAGACTGGAGACAGATCTGACGAGTACACACACTTCTGTAAACCGGCTTCGAGAGTT 672
QY               687 CTGCGTCCGCAAGAAGATGATCCGATGATCATCCGAGAGATGGAATATGTT 746
Db               673 CTGCGTCCGCAAGAAGATGATCCGATGATCATCCGAGAGATGGAATATGTT 732
QY               747 TGCATCAACCTGATGTGTCAGACATCCAGACATTTGATATGTGACATTTGCAATG 806
Db               733 GGCATCAACCTGATGTGATTAACACATCCAGACATCTTCAACGGGGACCTGTGCAAGCG 792
QY               807 GACCCGAGAGGGTCCAAATGGGGGAGCTCAAGCGGACCTTTTGTAGGACAGGAGCA 866
Db               793 GACCCGAGAGGGGCTCAATGGGGGAGCTCAAGCGGACCTTTTGTAGGACAGGAGCA 852
QY               867 CCTGGAGTCTGACCTTGGCAAGCGGTGACATTTGAGTCCAGAGAGAGCCCACTG 926
Db               853 CCTGGAGTCTTGGCAAGCGGTGACATTTGAGTCCAGAGAGAGCCCACTG 912
QY               927 AGGGGCTGC 935
Db               913 AGGACCAAGC 921

```

```

RESULT 15
AC078885/c       254993 bp   DNA      linear      HTG 15-MAY-2002
LOCUS            Mus musculus chromosome 6 clone RP23-117123 strain C57BL/6J,
DEFINITION       WORKING DRAFT SEQUENCE, 44 unordered pieces.
ACCESSION        AC078885
VERSION          AC078885.8 GI:18376842
KEYWORDS         HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE           Mus musculus (house mouse)
ORGANISM         Mus musculus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS          Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
                  Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Pereira, A.,
                  Gordon, M., Goltz, J.S. and Kucherlapati, R.
TITLE            High Throughput Mouse Sequencing
JOURNML          Unpublished
AUTHORS          2 (bases 1 to 254993)
                  Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
                  Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Pereira, A.,
                  Gordon, M., Goltz, J.S. and Kucherlapati, R.
TITLE            Direct Submissions
JOURNML          Submitted (08-NOV-2000) Department of Molecular Genetics, Albert
                  Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
                  Bronx, NY 10461, USA
                  On Jan 26, 2002 this sequence version replaced gi:14488282.
COMMENT
                  -----Genome Center
                  Center: Harvard Partners Genome Center
                  Center Code: HPGC
                  Web site: http://www.hpcgg.org/Sequence/mouse.html
                  Contact: hpcgcmendel.mgh.harvard.edu
                  -----Summary Statistics
                  Center Project name: AAF
                  Sequencing vector: pUC18; L08752
                  Chemistry: Dye-terminator Big Dye, 100%
                  *Consensus quality: 247096 at least 020
                  *Consensus quality: 242518 at least 030
                  *Consensus quality: 235589 at least 040

```



```
/note="assembly_name:Contig122"  
221185..222796  
/note="assembly_name:Contig121"  
222817..224762  
/note="assembly_name:Contig120"  
224783..225999  
/note="assembly_name:Contig119"  
226020..227409  
/note="assembly_name:Contig118"  
227430..230145  
/note="assembly_name:Contig117"  
230166..231842  
/note="assembly_name:Contig116"  
231863..232786  
/note="assembly_name:Contig115"  
232807..233137  
/note="assembly_name:Contig114"
```

Query Match 36.6%; Score 594.6; DB 2; Length 254993;
Best Local Similarity 79.8%; Pred.No.4.6e-115;
Matches 700; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

```
QY 148 CCGACCAAGCGCCCTTCTCTATAGGGTGAAGCGCGGCACTGCCAGCGGGAAGTCGACC 207  
DB 156715 CGGATCANNNNNNNNNNNNNNNNNNNNNNNNNGGGGGGGCGGCGACCGCTAGTGGCAAGTCAACA 156656  
QY 208 GTGTGTGAGAGATCATGTGAGTGTCTGGGACAGAACGAGTGGAAACAGCGGACGCGGAAG 267  
DB 156655 GTGTGTGAGAGATCATGTGAGTGTCTGGGACAGAACGAGTGGAAACAGCGGACGCGGAAG 156596  
QY 268 GTGTGTATCTGTAGCCAGGACGAGTGTCTACAAAGTCTGTACGCGACAGACAGAAAGGCCAAG 327  
DB 156595 TTGGTCAATCCTGAGCCAGAGATGTGCTTACAAAGTCTGTACGCGCGGACAGAAAGGCCAAG 156536  
QY 328 GCCTTGAAGAAGACGTACATTTTGAACCATCCAGATGCGCTTGAATATGATTTATGACAC 387  
DB 156535 GCTTTGAAGGACAGTACAAATTTTGAACCAACAGATGCTTTTGAATATGATTTATGACAC 156476  
QY 388 AGACTCTGAAGACATCTGTGAGGGCAAAACGGTGGAGTGCAGCCTATGATTTTGTG 447  
DB 156475 AAGACCTGGAAGAAATTTGTGAAAGCAAACTGTCCAGTCCCTACTATGATTTTGTG 156416  
QY 448 ACACACTCAGGTTACAGAGACCAACGGTGTCTACCTCGGACGTGTCTGTGTGAG 507  
DB 156415 ACCCACTCAGGTTACAGAGACCAACGTGTGTCTACCAAGTGTGTGTGTGTGTGAG 156356  
QY 508 GGCATCTTGTGTCTACAGCAGAGATCCGGACATGTTCCACCTCCGCTCTTCTGTG 567  
DB 156355 GGCATCTTGTGTCTACAGCAGAGATCCGGACATGTTCCACCTCCGCTCTTCTGTG 156296  
QY 568 GACACCGACTCGACGTCAGGCTGTCTGAAGATTCTCCGGACGTGCGCGGAGGAGG 627  
DB 156295 GACACCGACTCGATCTGATGTAGGCTGTCTGAAGATTCTCCGGAGTGTGCAACGAGGAGG 156236  
QY 628 GACCTGAGCAGATTCTGACGCAATACACCTTCTGTGAAGCGGCTTGTGAGAGTTC 687  
DB 156235 GACCTGAGAACGATCTGACTCACTATATACCGCTTGTGAAGAACGACCTTGTGAGAGTTC 156176  
QY 688 TGCCTGCCGACAAAGAGTATCCGATGTGATCATCCACGAGAGTGCACAAATATGTG 747  
DB 156175 TGCCTGCCGATTAAGAGTATCCGATGTGATCATCCCTGAGGAGTGTGAATATATGTG 156116  
QY 748 GCCATCAACTGATCTGTGACACATCCAGGACATTCATGATGTGACATCTGCAATGG 807  
DB 156115 GCCATCAACTGATCTGTGACACATCCAGGACATTCATGATGTGACATCTGTGCAAGAG 156056  
QY 808 CACCGAGAGGGTCCAAATGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACAC 867  
DB 156055 CACCGAGAGGGGCCCAACGGGCGCAACCAAGAGGACCTTCCCGAGCAGAGATCAC 155996  
QY 868 CTTGGGATGCTGACCTCTGGGAAAGGTCATTTTGAAGTCCAGACAGACCCCACTGA 927  
DB 155995 CTTGGGATGTTGGCACTGCAAGCGCTCACCTGGAGTCTTACAGACAGACCCCACTGA 155936
```

```
QY 928 GGGGCTCCGAGCCTCAGGSCAGATCTCCCGCCGCGATGTGTTCAGGGACTGAGCCT 987  
DB 155935 GAGCAGACAT---TGTAGTTCCCAACAGACCCAGATTTAAGGCGCTGGGATGACAC 155879  
QY 988 GGGGAGCGCCACCAACCAACCCACTGCTTCTCTGCGCG 1024  
DB 155878 TGGCTCTAAGACAGTCAAGGATCCCTCTCAGTG 155842
```

Search completed: November 25, 2003, 00:21:58
Job time : 6324.34 secs

QY 181 GGGGCACTGCCAGCGGGAAGTCGACCGCTGTGTGAGAAAGATCATGAGTTGCTGGACAG 240
 Db 181 GGGGCACTGCCAGCGGGAAGTCGACCGCTGTGTGAGAAAGATCATGAGTTGCTGGACAG 240
 QY 241 AACGAGGTGGAACAGCGGAGCGGGAAGTGGTCACTCTGAGCCAGAGAGTTCTACAG 300
 Db 241 AACGAGGTGGAACAGCGGAGCGGGAAGTGGTCACTCTGAGCCAGAGAGTTCTACAG 300
 QY 301 GTCTGACGCGAGAGCAAGAGGCCAAGGCTTTGAAAGACATGCAATTTTGAACATCCA 360
 Db 301 GTCTGACGCGAGAGCAAGAGGCCAAGGCTTTGAAAGACATGCAATTTTGAACATCCA 360
 QY 361 GATGCTTTGATATGATTTGATGACAGAGCTGTGAAGAACATCGGAGGGCAAAACG 420
 Db 361 GATGCTTTGATATGATTTGATGACAGAGCTGTGAAGAACATCGTGAAGGGCAAAACG 420
 QY 421 GTGAGAGTGCAGACTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGTGGTC 480
 Db 421 GTGAGAGTGCAGACTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGTGGTC 480
 QY 481 TACCTGCGGACGTGTTCTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGATCCGG 540
 Db 481 TACCTGCGGACGTGTTCTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGATCCGG 540
 QY 541 GACATGTTCCACCTGCGGCTTCTGTGAGACACGACCTCCAGAGCTGTCTCCAGAG 600
 Db 541 GACATGTTCCACCTGCGGCTTCTGTGAGACACGACCTCCAGAGCTGTCTCCAGAG 600
 QY 601 GTTCTCCGAGACGTGCCCGAGAGGAGACCTGAGAGAGATTTGACGACATGACCAACC 660
 Db 601 GTTCTCCGAGACGTGCCCGAGAGGAGACCTGAGAGAGATTTGACGACATGACCAACC 660
 QY 661 TTGCTGAAGCCGCTTCTGAGAGATTTGCTGCTGCGCAAAAGAAATGCCGATGTATC 720
 Db 661 TTGCTGAAGCCGCTTCTGAGAGATTTGCTGCTGCGCAAAAGAAATGCCGATGTATC 720
 QY 721 ATCCAGAGAGAGTGAACAATATGTTGCAATCAACTGATCGGAGACATCCAGAGAC 780
 Db 721 ATCCAGAGAGAGTGAACAATATGTTGCAATCAACTGATCGGAGACATCCAGAGAC 780
 QY 781 ATTCTGAATGGTGAATCTGCAATATGACACCGAGAGAGGTTCATATGGGCGAGACTACAG 840
 Db 781 ATTCTGAATGGTGAATCTGCAATATGACACCGAGAGAGGTTCATATGGGCGAGACTACAG 840
 QY 841 CGGACCTTTTCTGAGCCAGGGGACCACTCTGGGATGCTGACCTTGGCAAAACGCTCAT 900
 Db 841 CGGACCTTTTCTGAGCCAGGGGACCACTCTGGGATGCTGACCTTGGCAAAACGCTCAT 900
 QY 901 TTGGAGTCCAGACAGACCCCACTGAGGGGCTGCGAGGCTGAGGAGAGGTCTCCGCC 960
 Db 901 TTGGAGTCCAGACAGACCCCACTGAGGGGCTGCGAGGCTGAGGAGAGGTCTCCGCC 960
 QY 961 CGGATGTGTGTTGAGGAGCTGAGCTGAGGAGCGCCACCCACCACTGCTTCTCTC 1020
 Db 961 CGGATGTGTGTTGAGGAGCTGAGCTGAGGAGCGCCACCCACCACTGCTTCTCTC 1020
 QY 1021 GGGGCACTGAGAGAGTGTGAGAGAGGCTTCTCTCACTCACTGAGAGTGGAACTCAGA 1080
 Db 1021 GGGGCACTGAGAGAGTGTGAGAGAGGCTTCTCTCACTCACTGAGAGTGGAACTCAGA 1080
 QY 1081 TGTGTCACTGACTCAACTGTGAGGACACTGACAGGCGTCTGAGGTTTTCAGCCAC 1140
 Db 1081 TGTGTCACTGACTCAACTGTGAGGACACTGACAGGCGTCTGAGGTTTTCAGCCAC 1140
 QY 1141 TTAGGCTGTTGCGGTTTAAAGATCCCTTAGTCACTGAGAAATGACAGATGTGCA 1200
 Db 1141 TTAGGCTGTTGCGGTTTAAAGATCCCTTAGTCACTGAGAAATGACAGATGTGCA 1200
 QY 1201 GGAAGCTTGGAGAGCTTCTGTGAGGAAATGTGAGGACATTAATTTGGGAAATTTGAGAGAC 1260
 Db 1201 GGAAGCTTGGAGAGCTTCTGTGAGGAAATGTGAGGACATTAATTTGGGAAATTTGAGAGAC 1260

QY 1261 AGCTTAGACACTGGGCGGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGATT 1320
 Db 1261 AGCTTAGACACTGGGCGGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGATT 1320
 QY 1321 TTTTCCAGTGTGATCTGTTCTTACACTCACTCAACATTAATCAAAAGTTTGTGAACA 1380
 Db 1321 TTTTCCAGTGTGATCTGTTCTTACACTCACTCAACATTAATCAAAAGTTTGTGAACA 1380
 QY 1381 AGTACTTCTCTTTTACATGTTTACATGTCCTCAGTCTTCTGTTTCTGTTTCAATAA 1440
 Db 1381 AGTACTTCTCTTTTACATGTTTACATGTCCTCAGTCTTCTGTTTCTGTTTCAATAA 1440
 QY 1441 CAAGGCTGTTGAGGCTCAAACTAATTTCATGACCCAGTGTGAGTCCAGCCTG 1500
 Db 1441 CAAGGCTGTTGAGGCTCAAACTAATTTCATGACCCAGTGTGAGTCCAGCCTG 1500
 QY 1501 GCCTACAGGATATGAGGAGCACTGAGGAGATGTTTCCCTTGTGCTTTAAAG 1560
 Db 1501 GCCTACAGGATATGAGGAGCACTGAGGAGATGTTTCCCTTGTGCTTTAAAG 1560
 QY 1561 GCAGAGAACGAGGCGGATGCCCTGGAAGACACCCAGATCAACACAGCTTGTGGGG 1620
 Db 1561 GCAGAGAACGAGGCGGATGCCCTGGAAGACACCCAGATCAACACAGCTTGTGGGG 1620
 QY 1621 CCAG 1624
 Db 1621 CCAG 1624

RESULT 2
 US-09-833-381-2048
 ; Sequence 2048 Application US/09833381
 ; Patent No. US20020132090A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Robison, Keith E.
 ; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
 ; FILE REFERENCE: 5800-119
 ; CURRENT APPLICATION NUMBER: US/09/833,381
 ; PRIORITY FILING DATE: 2001-04-11
 ; PRIOR APPLICATION NUMBER: 09/516,448
 ; NUMBER OF SEQ ID NOS: 2050
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2048
 ; LENGTH: 1648
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(1648)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-381-2048

Query Match 86.9%; Score 1411; DB 10; Length 1648;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 46 GGGGAAAGGCGGCGCGGGGACCCGATGCGGAGCGGAGCGGAGATGCTTGGCG 105
 Db 36 GGGGAAAGGCGGCGCGGGGACCCGATGCGGAGCGGAGCGGAGATGCTTGGCG 95
 QY 106 GGAAGGAGAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCAAGGAGGCTTC 165
 Db 96 GGAAGGAGAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCAAGGAGGCTTC 155
 QY 166 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTCAACCGTGTGTGAGAAATCATG 225
 Db 156 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTCAACCGTGTGTGAGAAATCATG 215
 QY 226 GAGTTGCTGAGACAGAACGAGGTGAAACAGCGGACGGAAGGTGTCTATCTGAGCCAG 285
 Db 216 GAGTTGCTGAGACAGAACGAGGTGAAACAGCGGACGGAAGGTGTCTATCTGAGCCAG 275

QY 286 GACAGGTTCTACAGGTCCTGACGGCAGACAGAAAGCCAAAGCCTTGAAAGACAGTAC 345
| | | | |
Db 276 GACAGGTTCTACAGGTCCTGACGGCAGACAGAAAGCCAAAGCCTTGAAAGACAGTAC 335
| | | | |
QY 346 AATTTTGACCATTCAGATGCTTTGATATGATTTGATGACAGAGACTTGAAAGAACATC 405
| | | | |
Db 336 AATTTTGACCATTCAGATGCTTTGATATGATTTGATGACAGAGACTTGAAAGAACATC 395
| | | | |
QY 406 GTGAGAGGGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACACTCAAGTTACCA 465
| | | | |
Db 396 GTGAGAGGGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACACTCAAGTTACCA 455
| | | | |
QY 466 GAGACCAAGGTGCTACCTTGCGGACAGTGTCTGTTTGAAGGCACTTGATGTTCTAC 525
| | | | |
Db 456 GAGACCAAGGTGCTACCTTGCGGACAGTGTCTGTTTGAAGGCACTTGATGTTCTAC 515
| | | | |
QY 526 AGCCAGAGATCCGGGACATGTTTCAACTTGCGCTCTTGCTGACACCGACTCCGAGTTC 585
| | | | |
Db 516 AGCCAGAGATCCGGGACATGTTTCAACTTGCGCTCTTGCTGACACCGACTCCGAGTTC 575
| | | | |
QY 586 AGGCTGCTCGAAGAGTTCCTCCGGACAGTGGCCGAGAGGGAGACCTGAGACAGATTTCTG 645
| | | | |
Db 576 AGGCTGCTCGAAGAGTTCCTCCGGACAGTGGCCGAGAGGGAGACCTGAGACAGATTTCTG 635
| | | | |
QY 646 AGCAGATACACACCTTCGTGAAGCCGGCTTGAGAGAGTTCTGCCCGCAGAAAGAG 705
| | | | |
Db 636 AGCAGATACACACCTTCGTGAAGCCGGCTTGAGAGAGTTCTGCCCGCAGAAAGAG 695
| | | | |
QY 706 TA-TGCCGATGTGATCATCCACGAGAGTGTGACAAATATGTTGCCATCAACTGATTCGT 764
| | | | |
Db 696 TATTGCCGATGTGATCATCCACGAGAGTGTGACAAATATGTTGCCATCAACTGATTCGT 755
| | | | |
QY 765 GCAGCATCTCAGAGACATTTGAAATGTGTGACATCTGCAATTTGACACCGAGAGGGTCCAA 824
| | | | |
Db 756 GCAGCATCTCAGAGACATTTGAAATGTGTGACATCTGCAATTTGACACCGAGAGGGTCCAA 815
| | | | |
QY 825 TGAGCGAGGCTACAGAGGACCTTTTCTGAGCAGAGGGAGACACCTTGAGATGTCGACCTC 884
| | | | |
Db 816 TGAGCGAGGCTACAGAGGACCTTTTCTGAGCAGAGGGAGACACCTTGAGATGTCGACCTC 875
| | | | |
QY 885 TGAGCAACCGTCAATTTGAGAGTTCAGCAGACAGACCCCACTGAGGGGCTGCCAGCCTCA 944
| | | | |
Db 876 TGAGCAACCGTCAATTTGAGAGTTCAGCAGACAGACCCCACTGAGGGGCTGCCAGCCTCA 935
| | | | |
QY 945 GGGCAGAGTCTCCCGCCCGAGATGTGTGTTAGAGGACTGAGCCTGGGGAGCCCAACCA 1004
| | | | |
Db 936 GGGCAGAGTCTCCCGCCCGAGATGTGTGTTAGAGGACTGAGCCTGGGGAGCCCAACCA 995
| | | | |
QY 1005 CCCACTGCTTCTCTCGGCGCACCCCAAGGGAGTGTAGCAGGAGGCGCTTCTCACTCA 1064
| | | | |
Db 996 CCCACTGCTTCTCTCGGCGCACCCCAAGGGAGTGTAGCAGGAGGCGCTTCTCACTCA 1055
| | | | |
QY 1065 GAGGTGAAAACCTCAGATGTGTCTCACTCAGACTCACTTCTGGGACACTGACAGGCTTCC 1124
| | | | |
Db 1056 GAGGTGAAAACCTCAGATGTGTCTCACTCAGACTCACTTCTGGGACACTGACAGGCTTCC 1115
| | | | |
QY 1125 TGAAGTTTTCAGCAGCTTAGGCTCTGTTGCGGTTTAAAGATCCCTCTAGCTCAGTGA 1184
| | | | |
Db 1116 TGAAGTTTTCAGCAGCTTAGGCTCTGTTGCGGTTTAAAGATCCCTCTAGCTCAGTGA 1175
| | | | |
QY 1185 TGCCACAGATGTGACGAAAGCCTTGAGAGGCTTCTGAGGAGATGTAGGCACTATTTG 1244
| | | | |
Db 1176 TGCCACAGATGTGACGAAAGCCTTGAGAGGCTTCTGAGGAGATGTAGGCACTATTTG 1235
| | | | |
QY 1245 GGGAAAATTGAGAGACAGCTTAGACACTGGCTGGCCTGATGTTTGTGACATGTAACCC 1304
| | | | |
Db 1236 GGGAAAATTGAGAGACAGCTTAGACACTGGCTGGCCTGATGTTTGTGACATGTAACCC 1295
| | | | |
QY 1305 ACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTTACACACTCAGACATTAATCTC 1364
| | | | |
Db 1296 ACAGTGGAGAGAGTTTTCAGTCTGATCTGTTCTTTACACACTCAGACATTAATCTC 1355
| | | | |
QY 1365 AAAAGTTTGTGAAACAAGTACTTCTTTTATCATGTTACATGTTCTGATGTTTCTGT 1424
| | | | |

Db 1356 AAAAGTTTGTGAAACAAGTACTTCTTTTATCATGTTACATGTTCTCATGTTTCTGT 1415
| | | | |
QY 1425 TTTCTGTTTCATPACACAGAGCTGTTGTGGCTTACAAACCTTATTTCAATGACCACTGG 1484
| | | | |
Db 1416 TTTCTGTTTCATPACACAGAGCTGTTGTGGCTTACAAACCTTATTTCAATGACCACTGG 1475
| | | | |
QY 1485 TTTGCACTCAGCTGGGCGCTTACACGGAATATGGGGAGCCACTGAGGGATGTTTCCCGCT 1544
| | | | |
Db 1476 TTTGCACTCAGCTGGGCGCTTACACGGAATATGGGGAGCCACTGAGGGATGTTTCCCGCT 1535
| | | | |
QY 1545 TGCTTGTGCTTTAAAGGACAGAGAGCGCG 1576
| | | | |
Db 1536 TGCTTGTGCTTTAAAGGACAGAGAGCGCG 1567
| | | | |
RESULT 3
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xu, Aidong J.
; APPLICANT: Yang, Yonphong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungting
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-10-037-270-546
Query Match 57.4%; Score 932; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GGGGAAAGGGGCGGGGACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGGCG 105
| | | | |
Db 47 GGGGAAAGGGGCGGGGACCCGATGCGCGGAGCGGAGCGGAGATGCTTCGGCG 106
| | | | |
QY 106 GGAAGCGAAACTGGGAGAGCCCGCGCGAGAGCGGACCGTCCGCAACGAGGCGCTTC 165
| | | | |
Db 107 GGAAGCGAAACTGGGAGAGCCCGCGCGAGAGCGGACCGTCCGCAACGAGGCGCTTC 166
| | | | |
QY 166 CTGATAGGGGTAGAGCGCGGCACTGCAAGCGGAAATGCACTGTTGTGAGAAATGATG 225
| | | | |
Db 167 CTGATAGGGGTAGAGCGCGGCACTGCAAGCGGAAATGCACTGTTGTGAGAAATGATG 226
| | | | |

```
QY 226 GAGTTGCTGGGACAGAAAGAGTGGAAACAGCGGACAGCGGAAGGTGTCATCTTGAGCCAG 285
DB 227 GAGTTGCTGGGACAGAAAGAGTGGAAACAGCGGACAGCGGAAGGTGTCATCTTGAGCCAG 286
QY 286 GACAGGTTCTACAGAGTCTGACCGGACAGAGCAAGAGGCTTGAAAGGACAGTAC 345
DB 287 GACAGGTTCTACAGAGTCTGACCGGACAGAGCAAGAGGCTTGAAAGGACAGTAC 346
QY 346 AATTTTACCATCCAGATGCTTTTGAATATGATTTGATGCAAGAGACTTGAAAGATC 405
DB 347 AATTTTACCATCCAGATGCTTTTGAATATGATTTGATGCAAGAGACTTGAAAGATC 406
QY 406 GTGAGGAGCAAAAGGTGAGAGTGGCCAGACCTATATTTTGTGACACCTAAGTTTACA 465
DB 407 GTGAGGAGCAAAAGGTGAGAGTGGCCAGACCTATATTTTGTGACACCTAAGTTTACA 466
QY 466 GAGACCAAGGTGCTACACCTGCGGACGTGGTCTGTTTGAAGGACATCTTGTTTAC 525
DB 467 GAGACCAAGGTGCTACACCTGCGGACGTGGTCTGTTTGAAGGACATCTTGTTTAC 526
QY 526 AGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCACACCTCCGACGTC 585
DB 527 AGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCACACCTCCGACGTC 586
QY 586 AGGCTGCTGGAAGATGTTCCCGGGAAGTGGCCGAGAGGAGGACCTGAGACAGATTCG 645
DB 587 AGGCTGCTGGAAGATGTTCCCGGGAAGTGGCCGAGAGGAGGACCTGAGACAGATTCG 646
QY 646 ACAGAGTACACACCTTCTGTAAGCCCGGCTTCCAGAGAGTCTCTGCGGACAAAG 705
DB 647 ACAGAGTACACACCTTCTGTAAGCCCGGCTTCCAGAGAGTCTCTGCGGACAAAG 706
QY 706 TATGCCATGTATCATCCACGAGAGTGGACAAATATGTTGCTCATCACTGATGTG 765
DB 707 TATGCCATGTATCATCCACGAGAGTGGACAAATATGTTGCTCATCACTGATGTG 766
QY 766 CAGACATCCAGAGATCTGTAATGTTGACATCTGCAAAATGGACCCAGAGAGGCTCAAT 825
DB 767 CAGACATCCAGAGATCTGTAATGTTGACATCTGCAAAATGGACCCAGAGAGGCTCAAT 826
QY 826 GGGGAGAGTACAGCGGACCTTTCTGAGCCAGGAGACACCTTGGGATGCTGACCTCT 885
DB 827 GGGGAGAGTACAGCGGACCTTTCTGAGCCAGGAGACACCTTGGGATGCTGACCTCT 886
QY 886 GGCAGACGTTCACTTTGAGTCCAGACAGACCCCACTGAGGGGCTGCCAGCTCAG 945
DB 887 GGCAGACGTTCACTTTGAGTCCAGACAGACCCCACTGAGGGGCTGCCAGCTCAG 946
QY 946 GGCAGGTTCCCGGCGGCAATGTTTCAAG 977
DB 947 GGCAGGTTCCCGGCGGCAATGTTTCAAG 978

RESULT 4
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR APPLICATION DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804
```

```
Query Match 54.2%; Score 880; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 880; Conservative 0; Indels 0;

QY 745 GTTGCCATCAACCTGATCGTGCAGCATCCAGGACATTCGATATGATGACATCTGCAAA 804
DB 3901 GTTGCCATCAACCTGATCGTGCAGCATCCAGGACATTCGATATGATGACATCTGCAAA 3960
QY 805 TGGACCCGAGAGAGGTTCCAAATGGGCGGAGCTACAGAGGACCTTTTCTAGCCAGGGGAC 864
DB 3961 TGGACCCGAGAGAGGTTCCAAATGGGCGGAGCTACAGAGGACCTTTTCTAGCCAGGGGAC 4020
QY 865 CACCTGGGATGTGACCTCTGTCGCAAAAGGTCAATTTGAGTCCAGACAGACCCAC 924
DB 4021 CACCTGGGATGTGACCTCTGTCGCAAAAGGTCAATTTGAGTCCAGACAGACCCAC 4080
QY 925 TGAAGGAGCTGCCAGGCTCAGAGGAGTCTCCCGCCGACATGTGTTCAGGAGCTGAG 984
DB 4081 TGAAGGAGCTGCCAGGCTCAGAGGAGTCTCCCGCCGACATGTGTTCAGGAGCTGAG 4140
QY 985 CTTGGGAGAGCCCAACCCACACCACTGCTTCTCTGGGCGCACCCAGGAGGAGTTCAC 1044
DB 4141 CTTGGGAGAGCCCAACCCACACCACTGCTTCTCTGGGCGCACCCAGGAGGAGTTCAC 4200
QY 1045 AGCAGAGGCTTCTCTCACTCAGAGTGGAACTCAGATGTGTCACTCAGACTCAACTTGCT 1104
DB 4201 AGCAGAGGCTTCTCTCACTCAGAGTGGAACTCAGATGTGTCACTCAGACTCAACTTGCT 4260
QY 1105 GGGACACTGACAGAGGCTTCTGAGGTTTCAAGCACTTAAGGCTGTTGCGTTTAAAGAT 1164
DB 4261 GGGACACTGACAGAGGCTTCTGAGGTTTCAAGCACTTAAGGCTGTTGCGTTTAAAGAT 4320
QY 1165 CCTCTAGGTCACTGGAATATGCCACAGATGTGACAGAGGCTTGGAGGCTTCTGTGAG 1224
DB 4321 CCTCTAGGTCACTGGAATATGCCACAGATGTGACAGAGGCTTGGAGGCTTCTGTGAG 4380
QY 1225 GAATGTAGGACATTAATGGGAAATGAGAGACAGCTAGACACTGGCTGGCTGAT 1284
DB 4381 GAATGTAGGACATTAATGGGAAATGAGAGACAGCTAGACACTGGCTGGCTGAT 4440
QY 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 1344
DB 4441 GTTTTGTGACAGTGAACCCACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 4500
QY 1345 CACACTCACACATTAATCCAAAGTTTGTGAACAGATCTTCTTTTAAAGATTA 1404
DB 4501 CACACTCACACATTAATCCAAAGTTTGTGAACAGATCTTCTTTTAAAGATTA 4560
QY 1405 CATGCTCATGTTTCTGTTTCTGTTTCTGTTTCAATTAACAAGGCTGTTGTGCTTAACAAC 1464
DB 4561 CATGCTCATGTTTCTGTTTCTGTTTCTGTTTCAATTAACAAGGCTGTTGTGCTTAACAAC 4620
QY 1465 CTAAATTCATGACCCAGTGGTTTGACGTCAGGCTGCTTACAGGATATGGGAGCCAC 1524
DB 4621 CTAAATTCATGACCCAGTGGTTTGACGTCAGGCTGCTTACAGGATATGGGAGCCAC 4680
QY 1525 TGAAGGATGTTTCCCGGCTTGTGCTTAAAGGACAGAGGAGGAGGAGTCCCT 1584
DB 4681 TGAAGGATGTTTCCCGGCTTGTGCTTAAAGGACAGAGGAGGAGGAGTCCCT 4740
QY 1585 GGAAGACCCAGATCACACCCAGGCTTGTGCGGGGCCAG 1624
DB 4741 GGAAGACCCAGATCACACCCAGGCTTGTGCGGGGCCAG 4780

RESULT 5
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
```

```

; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match      54.2%; Score 880; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTCCATCAACTGATGTCGACGACATCCAGACATTTGTAATGTCATCTGCAA 804
DB 5997 GTTCCATCAACTGATGTCGACGACATCCAGACATTTGTAATGTCATCTGCAA 6056
QY 805 TGGCACCAGAGAGGTCCTCAATGGGCGGAGCTACAGGCGGACCTTTCTGAGCCAGGGGAC 864
DB 6057 TGGCACCAGAGAGGTCCTCAATGGGCGGAGCTACAGGCGGACCTTTCTGAGCCAGGGGAC 6116
QY 865 CACCTTGATGTCGACCTCTGGCAAAAGGTCACATTTGAGTCCAGACAGACCCAC 924
DB 6117 CACCTTGATGTCGACCTCTGGCAAAAGGTCACATTTGAGTCCAGACAGACCCAC 6176
QY 925 TGAAGGGCTCCGACGCTCAGGGAGGTCTCCGCCCGGACATGTGTTCAGGGACTGAG 984
DB 6177 TGAAGGGCTCCGACGCTCAGGGAGGTCTCCGCCCGGACATGTGTTCAGGGACTGAG 6236
QY 985 CCGGGGAGCGCCCAACCCAGCCGCTCTCTGGGCGCACCCAGGGGAGGTTCAG 1044
DB 6237 CCGGGGAGCGCCCAACCCAGCCGCTCTCTGGGCGCACCCAGGGGAGGTTCAG 6296
QY 1045 AGCGAGGCTCTCTCTCAGAGAGTGAATCTGAGTGTCTCAGACTCAACTTGTCT 1104
DB 6297 AGCGAGGCTCTCTCTCAGAGAGTGAATCTGAGTGTCTCAGACTCAACTTGTCT 6356
QY 1105 GGGACACTGACAGAGCGTTCTGAGAGTTTTCAGCCACTTACGCTGTGGCGTTAAAGAT 1164
DB 6357 GGGACACTGACAGAGCGTTCTGAGAGTTTTCAGCCACTTACGCTGTGGCGTTAAAGAT 6416
QY 1165 CCCTTAAGTCACTGAGAAATGCGACAGATGTCAGAGAGCTTGGAGAGCTTGTGAG 1224
DB 6417 CCCTTAAGTCACTGAGAAATGCGACAGATGTCAGAGAGCTTGGAGAGCTTGTGAG 6476
QY 1225 GAATGAGGACATTTTGGGAAATTTGAGAGACAGCTTACGCTGAGCTGAGCTGAT 1284
DB 6477 GAATGAGGACATTTTGGGAAATTTGAGAGACAGCTTACGCTGAGCTGAGCTGAT 6536
QY 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTTCTTA 1344
DB 6537 GTTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGGTTCTTA 6596
QY 1345 CACACTCAACAATACTCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 1404
DB 6597 CACACTCAACAATACTCAAAAGTTTGTGAACAAGTACTTCTTTTACATGTTA 6656
QY 1405 CATGCTCTCATGTTTCTGTTTCTGTTTCAATAACAAGAGTGTGTGGCTCAACAAC 1464
DB 6657 CATGCTCTCATGTTTCTGTTTCTGTTTCAATAACAAGAGTGTGTGGCTCAACAAC 6716
QY 1465 CTAAATTCATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACACGATATGGGAGCCAC 1524
DB 6717 CTAAATTCATGACCCAGTGGTTTTCAGTCCAGCGTGGCTTACACGATATGGGAGCCAC 6776
QY 1525 TGAAGGATGTTTTCCCGCTTGTGCTTGAAGGCAAGAGCGAGCGGATGCGCT 1584
DB 6777 TGAAGGATGTTTTCCCGCTTGTGCTTGAAGGCAAGAGCGAGCGGATGCGCT 6836
QY 1585 GGAAGCACCCAGCATCACCCAGGCTTGTGGGGGCGAG 1624
```

```

DB 6837 GGAAGCACCCAGCATCACCCAGGCTTGTGGGGGCGAG 6876

RESULT 6
US-09-896-522-3
; Sequence 3, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Gluckmann, Maria A.
; TITLE OF INVENTION: '57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-896-522-3

Query Match      51.4%; Score 834; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 ATGCTTCGCGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 153
DB 1 ATGCTTCGCGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGCGGCGGACCTGACGCGGAAAGTCAACCTGTGT 213
DB 61 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGCGGCGGACCTGACGCGGAAAGTCAACCTGTGT 120
QY 214 GAGAAGATCATGAGATGTTGCTGGGACAGACAGAGTGAACAGCGGACGCGGAAAGTGTGTC 273
DB 121 GAGAAGATCATGAGATGTTGCTGGGACAGACAGAGTGAACAGCGGACGCGGAAAGTGTGTC 180
QY 274 ATCTTAGGCGAGACAGGTTCTTCAAGAGTCTGACGCGAGAGCGAAGGCGGAGCTTGG 333
DB 181 ATCTTAGGCGAGACAGGTTCTTCAAGAGTCTGACGCGAGAGCGAAGGCGGAGCTTGG 240
QY 334 AAAGACATGATCAATTTTGAACCATCCAGATGCTTGTGATTAATGATTTGATGACAGAGCT 333
DB 241 AAAGACATGATCAATTTTGAACCATCCAGATGCTTGTGATTAATGATTTGATGACAGAGCT 300
QY 394 CTGAAGACATGCTGAGGCGAAGGCGGAGGAGTGCAGACCTATGATTTTGTGACACAC 453
DB 301 CTGAAGACATGCTGAGGCGAAGGCGGAGGAGTGCAGACCTATGATTTTGTGACACAC 360
QY 454 TCAAGGTTACAGAGACACAGGTGTCTTACCTGCGAGAGTGTCTGTTGAGGGCATC 513
DB 361 TCAAGGTTACAGAGACACAGGTGTCTTACCTGCGAGAGTGTCTGTTGAGGGCATC 420
QY 514 TTGCTGTTTACAGCCAGAGATCGGGACATGTTTCACTTGCGCTCTTCTGTGACACC 573
DB 421 TTGCTGTTTACAGCCAGAGATCGGGACATGTTTCACTTGCGCTCTTCTGTGACACC 480
QY 574 GACTCGAGGTACAGGTGCTGGAAGGTTCTCCGGAGAGTGTGCGGAGGAGGAGGAGCTG 633
DB 481 GACTCGAGGTACAGGTGCTGGAAGGTTCTCCGGAGAGTGTGCGGAGGAGGAGGAGCTG 540
QY 634 GAGCAGATTCGACGAGTACACCACTTTCGTGAAGCCCGGCTTTCGAGAGTTCGCTG 693
DB 541 GAGCAGATTCGACGAGTACACCACTTTCGTGAAGCCCGGCTTTCGAGAGTTCGCTG 600
QY 694 CCGACAAAGAGATATGCCATGTGATCATCCACAGAGAGTGAACAATATGTTGCGATC 753
DB 601 CCGACAAAGAGATATGCCATGTGATCATCCACAGAGAGTGAACAATATGTTGCGATC 660
QY 754 AACCTGATGTGAGACATCCAGGACATTTGAAATGTGACATCTGCAATGTGACACCGA 813
```

Db 661 AACCTGATCGTGCAGCATTCAGAGACTTGTGATGTGACATCTGCAATATGACACGA 720
Qy 814 GAGAGGCTTCATTTGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGACACCTTGGG 873
Db 721 GAGAGGCTTCATTTGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGACACCTTGGG 780
Qy 874 ATGCTGACCTCTGGCAAAAGGTGACATTTGAGTCCAGCAGACAGCCCACTGA 927
Db 781 ATGCTGACCTCTGGCAAAAGGTGACATTTGAGTCCAGCAGACAGCCCACTGA 834

RESULT 7

US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Rank, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g13899252, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183

Query Match 29.9%; Score 485; DB 12; Length 510;

Best Local Similarity 100.0%; Pred. No. 3.3e-250; Mismatches 0; Indels 0; Gaps 0;

Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 GTTSCATCAACCTGATCGTGCAGACATTCAGAGACTTGTGATGTGACATCTGCAAA 804
Db 26 GTTSCATCAACCTGATCGTGCAGACATTCAGAGACTTGTGATGTGACATCTGCAAA 85
Qy 805 TGGCAGCAGAGAGGCTCAATGGCGGAGCTAACAGCGACCTTTTCTGAGCCAGGGAC 864
Db 86 TGGCAGCAGAGAGGCTCAATGGCGGAGCTAACAGCGACCTTTTCTGAGCCAGGGAC 145
Qy 865 CACCTTGAGATGCTGACCTCTGGCAAAAGGTGACATTTGAGTCCAGACAGACCCAC 924
Db 146 CACCTTGAGATGCTGACCTCTGGCAAAAGGTGACATTTGAGTCCAGACAGACCCAC 205
Qy 925 TGAAGGGCTGCGAGGCTCAGGGAGGCTCCGCGCGGAGATGATGTTTCAAGGAGCTGAG 984
Db 206 TGAAGGGCTGCGAGGCTCAGGGAGGCTCCGCGCGGAGATGATGTTTCAAGGAGCTGAG 265
Qy 985 CTTGGGAGCGCCACCAACCACTGCTTCTCTGCGCGACCCAGGGAGGTGTAGC 1044
Db 266 CTTGGGAGCGCCACCAACCACTGCTTCTCTGCGCGACCCAGGGAGGTGTAGC 325
Qy 1045 AGCGAGGCTTCTCTCACTCAGAGATGGAATCAGATGTGTCACTGAGACTCAACTTGGT 1104
Db 326 AGCGAGGCTTCTCTCACTCAGAGATGGAATCAGATGTGTCACTGAGACTCAACTTGGT 385

Qy 1105 GGGACACTGACAGGCGGTTCTGTAGGTTTTCAGCCACTTAAGGCTGTTGGGTTTAAAGAT 1164
Db 386 GGGACACTGACAGGCGGTTCTGTAGGTTTTCAGCCACTTAAGGCTGTTGGGTTTAAAGAT 445
Qy 1165 CCCTTAGTCTACTGAGAAAATGCGACAGATGTGACAGAAAGCTTGGAGGCTTCTGTAG 1224
Db 446 CCCTTAGTCTACTGAGAAAATGCGACAGATGTGACAGAAAGCTTGGAGGCTTCTGTAG 505
Qy 1225 GAATG 1229
Db 506 GAATG 510

RESULT 8

US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-03-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379

Query Match 24.8%; Score 403; DB 11; Length 472;

Best Local Similarity 100.0%; Pred. No. 4.5e-206; Mismatches 0; Indels 0; Gaps 0;

Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCCAGCGGAGAGTGCACCGTGTGTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 249
Db 41 GCCAGCGGAGAGTGCACCGTGTGTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 100
Qy 250 GAAACGCGGAGCGGAGAGGTGTGATCTGAGCCAGAGAGTCTTCAAGGTCTGACG 309
Db 101 GAAACGCGGAGCGGAGAGGTGTGATCTGAGCCAGAGAGTCTTCAAGGTCTGACG 160
Qy 310 GCAGACAGAAAGGCCAAGGCTTTGAAAGAGCAGTAAATTTTGAACATCCAGATGCTTT 369
Db 161 GCAGACAGAAAGGCCAAGGCTTTGAAAGAGCAGTAAATTTTGAACATCCAGATGCTTT 220
Qy 370 GATATGATTTGATGACAGAGTCTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 429
Db 221 GATATGATTTGATGACAGAGTCTGAGAAATCATGATGATGCTGTGGACAGAAAGAGGTG 280
Qy 430 CCGACCTATGATTTTGTGACACTGAAAGTTTACAGAGACCAAGGTGTCTTCACTCTGCG 489
Db 281 CCGACCTATGATTTTGTGACACTGAAAGTTTACAGAGACCAAGGTGTCTTCACTCTGCG 340
Qy 490 GAGTGTGCTGTGTTTGAAGGACATCTTGTGTTTCAAGCCAGAGATCCGGGACATGTTG 549
Db 341 GAGTGTGCTGTGTTTGAAGGACATCTTGTGTTTCAAGCCAGAGATCCGGGACATGTTG 400
Qy 550 CACTGCGCTCTTCTGTGAGACCGGACTCGAGCTGAGGCTGT 592
Db 401 CACTGCGCTCTTCTGTGAGACCGGACTCGAGCTGAGGCTGT 443

RESULT 9

US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: A60MICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUATE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUATE 4.00e-26
; OTHER INFORMATION: NT HIT: g14783235, EVALUATE 1.00e-100
US-10-029-386-15883

Query Match 11.4%; Score 185; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.1e-88;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCCATCAACCTGATCGTGACGACATCCAGACATCTGTAAGTGTGACATCTGCAGAA 804
DB 3 GTTGCCATCAACCTGATCGTGACGACATCCAGACATCTGTAAGTGTGACATCTGCAGAA 62
QY 805 TGGCACCAGAGAGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGC 864
DB 63 TGGCACCAGAGAGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGC 122
QY 865 CACCTGGGATGCTGACCTCTGGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 924
DB 123 CACCTGGGATGCTGACCTCTGGGCAACGGTCACTTTGAGTCCAGACAGACCCAC 182
QY 925 TGAGG 929
DB 183 TGAGG 187

RESULT 10
US-09-764-877-399
; Sequence 399, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 222
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE

; LOCATION: (215)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-877-399

Query Match 3.6%; Score 58; DB 10; Length 222;
Best Local Similarity 100.0%; Pred. No. 2.4e-20;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1521 CCACGTGAGGAGTGTTCCTCCCTGCTGCTGCTTAAGGAGAGAGCGGCGGA 1578
DB 32 CCACGTGAGGAGTGTTCCTCCCTGCTGCTGCTTAAGGAGAGAGCGGCGGA 89

RESULT 11
US-09-908-975-24159
; Sequence 24159, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24159
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24159

Query Match 1.6%; Score 26; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 AGGCTGTCTCGAAGAGTTCTCCGGGA 611
DB 16 AGGCTGTCTCGAAGAGTTCTCCGGGA 41

RESULT 12
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Ruben
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220

Query Match 1.4%; Score 23; DB 10; Length 1310;

Best Local Similarity 100.0%; Pred. No. 0.16; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 ATCGTCAGCATCCAGACAT 782
 Db 891 ATCGTCAGCATCCAGACAT 913

RESULT 13

US-10-098-841-53
 ; Sequence 53, Application US/10098841
 ; Publication No. US20020197679A1
 ; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Xu, Chongjun
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Qian, Xiaohong B.
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20020197679A1e1. Nucleic Acids and
 ; FILE REFERENCE: 784CIP2
 ; CURRENT APPLICATION NUMBER: US/10/098,841
 ; PRIOR FILING DATE: 2002-03-13
 ; PRIOR APPLICATION NUMBER: 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 331
 ; SOFTWARE: pt_genes Version 1.0
 ; SEQ ID NO 53
 ; LENGTH: 1402
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (290)..(1075)
 ; US-10-098-841-53

Query Match 1.4%; Score 23; DB 13; Length 1402;

Best Local Similarity 100.0%; Pred. No. 0.16; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 ATCGTCAGCATCCAGACAT 782
 Db 950 ATCGTCAGCATCCAGACAT 972

RESULT 14

US-10-066-543-1843
 ; Sequence 1843, Application US/10066543
 ; Publication No. US20030087818A1
 ; GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu
 ; APPLICANT: Pyle, Ruth A.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Indrias, Carol Yoseph
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Fanger, Gary R.

; APPLICANT: Smith, Carole L.
 ; APPLICANT: Durham, Margarita
 ; APPLICANT: Scolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.563
 ; CURRENT APPLICATION NUMBER: US/10/066,543
 ; CURRENT FILING DATE: 2002-01-31
 ; NUMBER OF SEQ ID NOS: 3417
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1843
 ; LENGTH: 447
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-066-543-1843

Query Match 1.4%; Score 22; DB 14; Length 447;
 Best Local Similarity 100.0%; Pred. No. 0.57; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 AGGTGTCATCCAGACCGAGA 287
 Db 177 AGGTGTCATCCAGACCGAGA 198

RESULT 15

US-09-918-995-23923
 ; Sequence 23923, Application US/09918995
 ; Publication No. US20030073623A1
 ; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
 ; FILE REFERENCE: 20411-756
 ; CURRENT APPLICATION NUMBER: US/09/918,995
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: US/09/235,076
 ; PRIOR FILING DATE: 1999-01-20
 ; NUMBER OF SEQ ID NOS: 38054
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 23923
 ; LENGTH: 455
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(455)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-918-995-23923

Query Match 1.4%; Score 22; DB 11; Length 455;
 Best Local Similarity 100.0%; Pred. No. 0.57; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 AGGTGTCATCCAGACCGAGA 287
 Db 191 AGGTGTCATCCAGACCGAGA 212

Search completed: November 25, 2003, 07:26:58
 Job time : 1297.63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 02:04:12 ; Search time 113.64 Seconds
(Without alignments)
6307.673 Million cell updates/sec

Title: US-09-896-522-1

Sequence: 1 gggggggtcgcccgaccc.....ccagctctgctggggccag 1624

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 15

Total number of hits satisfying chosen parameters: 934

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	932	57.4	1288	4 US-09-620-312D-546	Sequence 546, App
2	783	48.2	834	4 US-09-536-647-1	Sequence 1, Appl
3	21	1.3	1529	3 US-08-858-876A-3	Sequence 3, Appl
4	21	1.3	1529	3 US-09-472-880-3	Sequence 3, Appl
5	21	1.3	3060	1 US-08-098-141-1	Sequence 1, Appl
6	19	1.2	1521	1 US-08-670-354-3	Sequence 3, Appl
7	19	1.2	1521	3 US-09-320-424-3	Sequence 3, Appl
8	19	1.2	1521	4 US-09-825-563-3	Sequence 3, Appl
9	19	1.2	1521	5 PCT-US96-10895-3	Sequence 3, Appl
10	19	1.2	1751	1 US-08-670-354-1	Sequence 1, Appl
11	19	1.2	1751	3 US-09-320-424-1	Sequence 1, Appl
12	19	1.2	1751	4 US-09-825-563-1	Sequence 1, Appl
13	19	1.2	1751	5 PCT-US96-10895-1	Sequence 1, Appl
14	19	1.2	1751	4 US-09-333-593A-5	Sequence 5, Appl
15	19	1.2	1769	4 US-09-505-350-1	Sequence 1, Appl
16	19	1.2	1769	4 US-09-016-434-1072	Sequence 1072, Ap
17	19	1.2	1230025	4 US-09-198-452A-1	Sequence 1, Appl
18	18	1.1	1338	4 US-09-252-991A-6897	Sequence 6897, Ap
19	18	1.1	1338	4 US-09-252-991A-6934	Sequence 6934, Ap
20	18	1.1	2726	4 US-09-252-991A-6861	Sequence 6861, Ap
21	18	1.1	3127	4 US-08-461-823-1	Sequence 1, Appl
22	18	1.1	3127	4 US-09-620-312D-613	Sequence 613, App
23	18	1.1	4264	2 US-08-784-649A-1	Sequence 1, Appl
24	18	1.1	4264	2 US-08-784-649A-5	Sequence 5, Appl
25	18	1.1	4646	1 US-08-181-471-2	Sequence 2, Appl
26	18	1.1	4669	2 US-08-583-276-18	Sequence 18, Appl
27	18	1.1	4669	2 US-08-752-447-1	Sequence 1, Appl

28	18	1.1	4669	4 US-09-316-167-1	Sequence 1, Appl
29	18	1.1	4669	6 5206352-3	Patent No. 5206352
30	18	1.1	6505	2 US-08-793-610-5	Sequence 5, Appl
31	18	1.1	8630	4 US-09-306-417-1	Sequence 1, Appl
32	18	1.1	8630	4 US-09-306-417-2	Sequence 2, Appl
33	18	1.1	9318	2 US-08-793-610-6	Sequence 6, Appl
34	17	1.0	245	2 US-08-691-814B-76	Sequence 76, Appl
35	17	1.0	275	2 US-08-691-814B-79	Sequence 79, Appl
36	17	1.0	276	2 US-08-691-814B-78	Sequence 78, Appl
37	17	1.0	290	2 US-08-691-814B-80	Sequence 80, Appl
38	17	1.0	300	2 US-08-691-814B-75	Sequence 75, Appl
39	17	1.0	383	2 US-08-691-814B-74	Sequence 74, Appl
40	17	1.0	431	2 US-08-691-814B-84	Sequence 84, Appl
41	17	1.0	543	2 US-08-691-814B-81	Sequence 81, Appl
42	17	1.0	592	3 US-09-385-982-265	Sequence 265, App
43	17	1.0	600	4 US-09-252-991A-4602	Sequence 4602, Ap
44	17	1.0	672	4 US-09-252-991A-15543	Sequence 15543, A
45	17	1.0	749	1 US-08-579-667-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-620-312D-546
; Sequence 546, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jiao-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungting
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilui
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_ftl_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-09-620-312D-546

Query Match 57.4%; Score 932; DB 4; Length 1288;
Best Local Similarity 100.0%; Pred No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGGAAGGGCGGGCGGGAACCGATGCGGGAGCGAGCGGAGATGCTTCGGCG 105
DB 47 GGGGAAGGGCGGGCGGGAACCGATGCGGGAGCGAGCGGAGATGCTTCGGCG 106

QY 106 GAGAGCGAAGACTGCGAGAGCCCGCGCCGAGAGCCGACCGTCCGACACAGCGGCCCTTC 165
DB 107 GGAAGCGAAGACTGCGAGAGCCCGCGCCGAGAGCCGACCGTCCGACACAGCGGCCCTTC 166
QY 166 CTGATAGAGGGGAGACGGCGGAGCTGCGAGCGGAAAGTGCACCGTGTGAGAAATCATG 225
DB 167 CTGATAGAGGGGAGACGGCGGAGCTGCGAGCGGAAAGTGCACCGTGTGAGAAATCATG 226
QY 226 GAGTGTCTGGAGCAGAAAGAGGTGAAACAGCGGACCGGAAAGTGTGATCTTGAGCCAG 285
DB 227 GAGTGTCTGGAGCAGAAAGAGGTGAAACAGCGGACCGGAAAGTGTGATCTTGAGCCAG 286
QY 286 GAGAGGTTCTAAGAGTCTGAGCGGAGAGAAAGCGCTTGAAGAGCACTAC 345
DB 287 GAGAGGTTCTAAGAGTCTGAGCGGAGAGAAAGCGCTTGAAGAGCACTAC 346
QY 346 AATTTTACATCCAGATGCTTGTATGATTTGATGACAGAGACTTGAAGAAATC 405
DB 347 AATTTTACATCCAGATGCTTGTATGATTTGATGACAGAGACTTGAAGAAATC 406
QY 406 GTGAGAGCGAAGAGGTGAGAGTCCGACCTATGATTTTGTGACACTCAAGGTTACA 465
DB 407 GTGAGAGCGAAGAGGTGAGAGTCCGACCTATGATTTTGTGACACTCAAGGTTACA 466
QY 466 GAGAGCAGGAGTCTACCGTCCGAGAGTGTGATGAGGAGCTTGTGTCTAC 525
DB 467 GAGAGCAGGAGTCTACCGTCCGAGAGTGTGATGAGGAGCTTGTGTCTAC 526
QY 526 AGCAGAGAGATCCGAGACATGTTCCACTGCGCTCTGTCGACACCGACTCCGAGCTC 585
DB 527 AGCAGAGAGATCCGAGACATGTTCCACTGCGCTCTGTCGACACCGACTCCGAGCTC 586
QY 586 AGGCTGTCTCGAAGATTTCTCCGGAAGTGTGCGGAGGAGAGCTGAGACAGATTTCTG 645
DB 587 AGGCTGTCTCGAAGATTTCTCCGGAAGTGTGCGGAGGAGAGCTGAGACAGATTTCTG 646
QY 646 AGCAGAGACACACCTTCTGTAAGCGGCGCTTGAAGAGTCTGCTGCGGACAAAGAG 705
DB 647 AGCAGAGACACACCTTCTGTAAGCGGCGCTTGAAGAGTCTGCTGCGGACAAAGAG 706
QY 706 TATCCGATGATCATCCACAGAGAGTGAACAATATGTTGCATCACTGATGCTG 765
DB 707 TATCCGATGATCATCCACAGAGAGTGAACAATATGTTGCATCACTGATGCTG 766
QY 766 CAGACATCCAGAGATTTCTGAATGTGACATCTTGCAATGTGACCGAGAGGTTCAAT 825
DB 767 CAGACATCCAGAGATTTCTGAATGTGACATCTTGCAATGTGACCGAGAGGTTCAAT 826
QY 826 GGGGAGAGCTCAAGCGGAGCTTCTTGAAGCAGGAGGACACCTGAGGATGCTGACCT 885
DB 827 GGGGAGAGCTCAAGCGGAGCTTCTTGAAGCAGGAGGACACCTGAGGATGCTGACCT 886
QY 886 GGCAGAGCTGACATTTGAGATCCAGAGAGACCCCACTGAGGAGCTGCGGAGCTCAG 945
DB 887 GGCAGAGCTGACATTTGAGATCCAGAGAGACCCCACTGAGGAGCTGCGGAGCTCAG 946
QY 946 GGCAGGTTCTCCCGCCGAGATGTGTTCAAG 977
DB 947 GGCAGGTTCTCCCGCCGAGATGTGTTCAAG 978

RESULT 2
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Johnson, Randall
; APPLICANT: Ho, Yen Sen
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536, 647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-09-536-647-1
Query Match 48.2%; Score 783; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGCTTCCGCGGAGGCGAAGCTGCGAGAGCTCCCGCCGAGAGCCGACCGTCCGAC 153
DB 1 ATGCTTCCGCGGAGGCGAAGCTGCGAGAGCTCCCGCCGAGAGCCGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGACCTGCGAGCGGAAAGTGCACCTGTGT 213
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGACCTGCGAGCGGAAAGTGCACCTGTGT 120
QY 214 GAGAGATCATGAGTGTGCGGACAGAAAGAGTGAACAGCGGACGCGAAGTGTCT 273
DB 121 GAGAGATCATGAGTGTGCGGACAGAAAGAGTGAACAGCGGACGCGAAGTGTCT 180
QY 274 ATCTGAGCCAGAGACAGTGTCTCAAGGTCTGACGCGAGAGAAAGCCAAAGCTTG 333
DB 181 ATCTGAGCCAGAGACAGTGTCTCAAGGTCTGACGCGAGAGAAAGCCAAAGCTTG 240
QY 334 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTATGATTTGATGACAGAGCT 393
DB 241 AAAGGACAGTCAATTTTGAACATCCAGATGCTTGTATGATTTGATGACAGAGCT 300
QY 394 CTGAAGACATCTGAGAGGCGAAGAGTGTGAGAGTCCGACCTATGATTTTGTGACAC 453
DB 301 CTGAAGACATCTGAGAGGCGAAGAGTGTGAGAGTCCGACCTATGATTTTGTGACAC 360
QY 454 TCAAGTTACAGAGACACAGTGTCTACCTGCGGAGTGTCTGTTGAGGAGCTC 513
DB 361 TCAAGTTACAGAGACACAGTGTCTACCTGCGGAGTGTCTGTTGAGGAGCTC 420
QY 514 TTGAGTTCAAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTCAC 573
DB 421 TTGAGTTCAAGCCAGAGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTCAC 480
QY 574 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCGGAGAGTGTGCGGAGAGAGCTG 613
DB 481 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCGGAGAGTGTGCGGAGAGAGCTG 540
QY 613 GAGCAGTTTGAAGCAGATACACACCTTGTGAAGCCGCGCTTCAAGAGATTTCTGCT 693
DB 541 GAGCAGTTTGAAGCAGATACACACCTTGTGAAGCCGCGCTTCAAGAGATTTCTGCT 600
QY 694 CCGAGAAAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGTCATC 753
DB 601 CCGAGAAAGAGTATGCGGATGATCATCCAGAGAGTGAACAATATGTTGTCATC 660
QY 754 AACCTGATGTCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 813
DB 661 AACCTGATGTCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 720
QY 814 GAGAGTCCAAATGGGAGGAGTCAAGCGGACCTTTTCTGAGCAGGAGACCACTGTGG 873
DB 721 GAGAGTCCAAATGGGAGGAGTCAAGCGGACCTTTTCTGAGCAGGAGACCACTGTGG 780
QY 874 ATGCTGACCTGAGCAAGCGTCAATTTGAGATCCAGAGAGAGCCCACTGA 927
DB 781 ATGCTGACCTGAGCAAGCGTCAATTTGAGATCCAGAGAGAGCCCACTGA 834

RESULT 3
US-08-858-876A-3/C

; Sequence 3, Application US/08858876A
; Patent No. 6022856
; GENERAL INFORMATION:
; APPLICANT: Daniel CAPUT
; APPLICANT: Pascale CHALON
; APPLICANT: Pascale FERRARA
; APPLICANT: Vita NATALIO
; TITLE OF INVENTION: Type 2 Neurotensin Receptor
; TITLE OF INVENTION: (hnt-R2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,876A
; FILING DATE: 19-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1529 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41..1288
; US-08-858-876A-3
; Query Match 1.3%; Score 21; DB 3; Length 1529;
; Best Local Similarity 100.0%; Pred. No. 0.86;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; QY 1048 GAGGCTTCTCTCAGAG 1068
; DB 826 GAGGCTTCTCTCAGAG 806
; RESULT 4
; US-09-472-880-3/C
; Sequence 3, Application US/09472880
; Patent No. 6274333
; GENERAL INFORMATION:
; APPLICANT: Daniel CAPUT
; APPLICANT: Pascale CHALON
; APPLICANT: Pascale FERRARA
; APPLICANT: Vita NATALIO
; TITLE OF INVENTION: Type 2 Neurotensin Receptor
; TITLE OF INVENTION: (hnt-R2)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/472,880
; FILING DATE: 28-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1529 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 41..1288
; US-09-472-880-3
; Query Match 1.3%; Score 21; DB 3; Length 1529;
; Best Local Similarity 100.0%; Pred. No. 0.86;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; QY 1048 GAGGCTTCTCTCAGAG 1068
; DB 826 GAGGCTTCTCTCAGAG 806
; RESULT 5
; US-08-098-141-1
; Sequence 1, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/098,141
; FILING DATE: 19930723
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan Beg., Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: 092662-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617:227-5020
; TELEFAX: 617:227-7566
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

Query Match 1.3%; Score 21; DB 1; Length 3060;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 272 TCATCTGAGCCAGACAGGT 292
DB 1344 TCATCTGAGCCAGACAGGT 1364

RESULT 6
US-08-670-354-3/c
Sequence 3, Application US/08670354
Patent No. 5763223
GENERAL INFORMATION:
APPLICANT: Steven R. Wiley and
APPLICANT: Raymond G. Goodwin.
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,354
FILING DATE: 25-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: HuA1C-dv
FEATURE:
NAME/KEY: CDS
LOCATION: 78..383

US-08-670-354-3

Query Match 1.2%; Score 19; DB 1; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1415 TGTTTCTGTTTCTGTTT 1433
DB 891 TGTTTCTGTTTCTGTTT 873

RESULT 7
US-09-320-424-3/c
Sequence 3, Application US/09320424
Patent No. 6284236
GENERAL INFORMATION:
APPLICANT: Wiley, Steven R.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Cytokine that Induces Apoptosis
FILE REFERENCE: 2835-E
CURRENT APPLICATION NUMBER: US/09/320,424
CURRENT FILING DATE: 1999-05-26
EARLIER APPLICATION NUMBER: 09/190,046
EARLIER FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 09/048,641
EARLIER FILING DATE: 1998-03-26
EARLIER APPLICATION NUMBER: 08/670,354
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 08/548,368
EARLIER FILING DATE: 1995-11-01
EARLIER APPLICATION NUMBER: 08/496,632
EARLIER FILING DATE: 1995-06-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 1521
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (78)..(383)
US-09-320-424-3

Query Match 1.2%; Score 19; DB 3; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1415 TGTTTCTGTTTCTGTTT 1433
DB 891 TGTTTCTGTTTCTGTTT 873

RESULT 8
US-09-825-563-3/c
Sequence 3, Application US/09825563
Patent No. 6521228
GENERAL INFORMATION:
APPLICANT: Wiley, Steven R.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Cytokine that Induces Apoptosis
FILE REFERENCE: 2835-E
CURRENT APPLICATION NUMBER: US/09/825,563
CURRENT FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 09/320,424
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: 09/190,046
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 09/048,641
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 08/670,354
PRIOR FILING DATE: 1996-06-25
PRIOR APPLICATION NUMBER: 08/548,368
PRIOR FILING DATE: 1995-11-01

PRIOR APPLICATION NUMBER: 08/496,632
PRIOR FILING DATE: 1995-06-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 1521
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (78) .. (383)
US-09-825-563-3

Query Match 1.2%; Score 19; DB 4; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 TGTTCCTGTTTCGTTT 1433
Db 891 TGTTCCTGTTTCGTTT 873

RESULT 9
PCT-US96-10895-3/c
Sequence 3, Application PC/TUS9610895
GENERAL INFORMATION:
APPLICANT: Immunex Corporation.
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10895
FILING DATE: 25-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
MOLTYPE: cDNA to mRNA
FEATURE:
NAME/KEY: NO
IMMEDIATE SOURCE:
CLONE: HuAIC-dv
FEATURE:

NAME/KEY: CDS
LOCATION: 78..383
PCT-US96-10895-3

Query Match 1.2%; Score 19; DB 5; Length 1521;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1415 TGTTCCTGTTTCGTTT 1433
Db 891 TGTTCCTGTTTCGTTT 873

RESULT 10
US-08-670-354-1/c
Sequence 1, Application US/08670354
Patent No. 5763223
GENERAL INFORMATION:
APPLICANT: Steven R. Wiley and
Raymond G. Goodwin
TITLE OF INVENTION: Cytokine That Induces Apoptosis
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple 7.5.2
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/670,354
FILING DATE: 25-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,632
FILING DATE: 29-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/548,368
FILING DATE: 01-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2835-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756822
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOTHETICAL: NO
MOLTYPE: cDNA to mRNA
FEATURE:
NAME/KEY: NO
IMMEDIATE SOURCE:
CLONE: huAIC
FEATURE:
NAME/KEY: CDS
LOCATION: 88..933
US-08-670-354-1

Query Match 1.2%; Score 19; DB 1; Length 1751;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

This Page Blank (uspto)

PN WO200202761-A2

DNA of a human uro-
 Human poly(nucleoti-
 Human cDNA clone (c-
 Human uridine kinase
 Human mdt cDNA in
 Human musculoskele-
 cDNA encoding novel
 Mouse spliced trans-
 Human gene express-
 Human prostate can-
 Human transferrase
 Human ovarian anti-
 Human poly(nucleoti-
 Human prostate exp-
 Human prostate exp-
 Human prostate exp-
 Human prostate exp-
 Human prostate exp-
 Human prostate exp-
 Human gene express-
 Anticancer agent t
 Oligonucleotide fo
 Oligonucleotide fo
 Oligonucleotide fo
 DNA encoding novel
 Human poly(nucleoti-
 Human uridine kinase
 Urea transporter p
 Human uridine kinase
 Human uridine kinase
 Oligonucleotide fo
 Oligonucleotide fo
 DNA encoding novel
 Human bone marrow
 DNA encoding novel
 Drosohilla melanog
 Human Her-2/neu ov
 Drosohilla melanog

CC The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as haematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing,
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.

Query Match	100.0%	Score 1624;	DB 24;	Length 1624;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1624; Conservative	0;	Mismatches	0;	Gaps 0;

OY 1 GTGGGGTCCGCTTCGCACTTCGCGCTTGGGCGGGGCGCGCGGGGCCCGGGGGAAGGGGCGGGC 60
 |||||
 |||||
 Db 1 GTGGGGTCCGCTTCGCACTTCGCGCTTGGGCGGGGCGCGCGGGGCCCGGGGGAAGGGGCGGGC 60
 |||||
 |||||
 OY 61 GCGGGGACCCGATGCGCGGGGAGCGGAGGCCGAATGGCTTTGGCGGGAGGCGAAAGACTGC 120
 |||||
 |||||
 Db 61 GCGGGGACCCGATGCGCGGGGAGCGGAGGCCGAATGGCTTTGGCGGGAGGCGAAAGACTGC 120
 |||||
 |||||
 OY 121 GAGAGCCCGCGCGCGGAGGCCACGTCGCGCACAGCGGCCCTTCTCGATAGGGGTGAGC 180
 |||||
 |||||
 Db 121 GAGAGCCCGCGCGCGGAGGCCACGTCGCGCACAGCGGCCCTTCTCGATAGGGGTGAGC 180
 |||||
 |||||
 OY 181 GCGGCACTGCCAGCGGGGAAGTGACCGTGTGAGAGAATCATGAGTTGGGCGAG 240
 |||||
 |||||
 Db 181 GCGGCACTGCCAGCGGGGAAGTGACCGTGTGAGAGAATCATGAGTTGGGCGAG 240
 |||||
 |||||
 OY 241 AACGAGGTGAAACAGGGGCGAGGGAGGTGTATCCTGAGCCAGGACAGTTCTAACG 300
 |||||
 |||||
 Db 241 AACGAGGTGAAACAGGGGCGAGGGAGGTGTATCCTGAGCCAGGACAGTTCTAACG 300
 |||||
 |||||
 OY 301 GTCTCTGACGCGAGAGCAAGGCGCAAGGCTTGAAAGGACAGTAACTTTGACCATCCA 360
 |||||
 |||||
 Db 301 GTCTCTGACGCGAGAGCAAGGCGCAAGGCTTGAAAGGACAGTAACTTTGACCATCCA 360
 |||||
 |||||
 OY 361 GATGACCTTTGATATGATTTGATGCAACAGACTCTGAAGAACATCGTGAGGGCAAAACG 420
 |||||
 |||||
 Db 361 GATGACCTTTGATATGATTTGATGCAACAGACTCTGAAGAACATCGTGAGGGCAAAACG 420
 |||||
 |||||
 OY 421 GTGGAGTCCGCACTATGATTTTGTGCACACTCAAGGTTTACAGAGACCAAGGTGTC 480

Db	421	GTGAGAGTGCAGACTTATGATTTTGTGACACACTCAAGTTTACCAGAGACCAAGTGTGTC	480
Qy	481	TACCTCGCGAGCGTGTCTGTGTTTGAAGGGACCTTGTGTCTTACACGACCAAGATCCGG	540
Db	481	TACCTCGCGAGCGTGTCTGTGTTGAAGGGACCTTGTGTCTTACACGACCAAGATCCGG	540
Qy	541	GACATGTTTCCACTGTGCGCTCTTGTGTGACACCGACTTCCGACCGTCAAGCTGTCTCGAAGA	600
Db	541	GACATGTTTCCACTGTGCGCTCTTGTGTGACACCGACTTCCGACCGTCAAGCTGTCTCGAAGA	600
Qy	601	GTTCTCCGGGAGCGTGCAGCGGAGGAGGAGCCTGAGACAGATTCTGACGACGATACCAACC	660
Db	601	GTTCTCCGGGAGCGTGCAGCGGAGGAGGAGCCTGAGACGAGATTCTGACGACGATACCAACC	660
Qy	661	TTTGTGAAGCCGAGCTTTCAGAGAGTTTGTGCTGCTGCGCAAAAGATATGCCGATGTATC	720
Db	661	TTTGTGAAGCCGAGCTTTCAGAGAGTTTGTGCTGCTGCGCAAAAGATATGCCGATGTATC	720
Qy	721	ATCCACGAGAGAGTGTGACATATGTGTGCCATCAACTGATGTGTGACGACATCTCAGAC	780
Db	721	ATCCACGAGAGAGTGTGACATATGTGTGCCATCAACTGATGTGTGACGACATCTCAGAC	780
Qy	781	ATTCTGAATGTGTGACATCTGTCAAAATGTGCACCGAGAGAGGTCCAAATGGGCGGAGCTACAG	840
Db	781	ATTCTGAATGTGTGACATCTGTCAAAATGTGCACCGAGAGAGGTCCAAATGGGCGGAGCTACAG	840
Qy	841	CGGACCTTTTCTGAGCCAGAGGGAGCACCCCTGTGGATGTGTGACCTTGTGCAACCGTTCACAT	900
Db	841	CGGACCTTTTCTGAGCCAGAGGGAGCACCCCTGTGGATGTGTGACCTTGTGCAACCGTTCACAT	900
Qy	901	TTTGAAGTTCAGACAGACAGCCCACTGAGGGGCTGTCCAGAGCTCAGGGCAGGTCTTCCGGC	960
Db	901	TTTGAAGTTCAGACAGACAGCCCACTGAGGGGCTGTCCAGAGCTCAGGGCAGGTCTTCCGGC	960
Qy	961	CGGATGTGTGTTTGAAGGACTGAGCCGTGGGAGGCGCCACCACTGAGCTTCTCTC	1020
Db	961	CGGATGTGTGTTTGAAGGACTGAGCCGTGGGAGGCGCCACCACTGAGCTTCTCTC	1020
Qy	1021	GGCCACCCCGAGGGAGTGTGAGAGGAGGCGCTTCTCTCACTCAGAGTGGAAACTCAGA	1080
Db	1021	GGCCACCCCGAGGGAGTGTGAGAGGAGGCGCTTCTCTCACTCAGAGTGGAAACTCAGA	1080
Qy	1081	TGTGTCACTCAGACTCAACTGTGCTGGGACACTGACAGGCGTTCTGTGAGTTTCAAGCCAC	1140
Db	1081	TGTGTCACTCAGACTCAACTGTGCTGGGACACTGACAGGCGTTCTGTGAGTTTCAAGCCAC	1140
Qy	1141	TTAAGGCTGTGTGGGTTTAAAGATCCCTCTAAGTCACTGAGAAATGCCACAGAAATGTGCA	1200
Db	1141	TTAAGGCTGTGTGGGTTTAAAGATCCCTCTAAGTCACTGAGAAATGCCACAGAAATGTGCA	1200
Qy	1201	GGAAGCCTGGGAGGCTTCTGTGAGGAATGTGAGGACATTAATGGGGAAATTGAGAGAC	1260
Db	1201	GGAAGCCTGGGAGGCTTCTGTGAGGAATGTGAGGACATTAATGGGGAAATTGAGAGAC	1260
Qy	1261	AGCCTAAGCACTGGCTGCGCTGATGTTTGTGTGACAGTGAACCCACAGTGGAGAGATT	1320
Db	1261	AGCCTAAGCACTGGCTGCGCTGATGTTTGTGTGACAGTGAACCCACAGTGGAGAGATT	1320
Qy	1321	TTTTTCAGTCTGATCTGGTCTTTCACACTCAGACATTAACCTCMAAAGTTTGTGTGACA	1380
Db	1321	TTTTTCAGTCTGATCTGGTCTTTCACACTCAGACATTAACCTCMAAAGTTTGTGTGACA	1380
Qy	1381	AGTACTTTCCTTTTTCACATGTACATGTCTCAATGTTTCTGTCTTCTGTTCATTAACA	1440
Db	1381	AGTACTTTCCTTTTTCACATGTACATGTCTCAATGTTTCTGTCTTCTGTTCATTAACA	1440
Qy	1441	CAAGGCTGTGTGTGGCTTACAAACCTAATTTCAATGACCCAGTGTGTTCAGTCCAGCGTG	1500
Db	1441	CAAGGCTGTGTGTGGCTTACAAACCTAATTTCAATGACCCAGTGTGTTCAGTCCAGCGTG	1500
Qy	1501	GCCTAACGATATAGGGAGCCACTGAGGAGTGTTTTCCCTTGTCTGTGTCTTAAAG	1560

DB 1501 GCCTACCGATATGAGGAGCCACTGAGGATGTTTCCCTGCTTGCTTAAG 1560
QY 1561 GCAGAGAGCGAGCGCGATGCTTGGAGAGCCAGATACGCCAGGTTGTGGGG 1620
DB 1561 GCAGAGAGCGAGCGCGATGCTTGGAGAGCCAGATACGCCAGGTTGTGGGG 1620
QY 1621 CCAG 1624
DB 1621 CCAG 1624

RESULT 2
AA160444
ID AA160444 standard; cDNA; 2152 BP.
AC AA160444;
XX
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 4433.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
XX Homo sapiens.
XX
XX MO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-0552317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX P-PSDB; AAM41288.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Claim 1; SEQ ID NO 4433; 10078BP; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening
XX assays for receptor activity, arthritis and inflammation, leukemias and
XX C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Query Match 94.1%; Score 1528; DB 22; Length 2152;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 46 GGGGAAAGGGCGCGCGCGCGGAGCCGATGCGCGGAGCCGAGCCGAGATGCTTGGCG 105
DB 11 GGGGAAAGGGCGCGCGCGCGGAGCCGATGCGCGGAGCCGAGATGCTTGGCGG 70
QY 106 GGAAGCGGAAGTGTGAGAGCCCGCGCGGAGCCGCGGAGCCGAGATGCTTGGCG 165
DB 71 GGAAGCGGAAGTGTGAGAGCCCGCGCGGAGCCGCGGAGCCGAGATGCTTGGCG 130
QY 166 CTGATAGGGGTGAGGCGCGGCACTGCGAGCGGGAAGTCAGCCGTGTGAGAGATCATG 225
DB 131 CTGATAGGGGTGAGGCGCGGCACTGCGAGCGGGAAGTCAGCCGTGTGAGAGATCATG 190
QY 226 GAGTTGCTGGGACAGACGAGGTGAAACGCGGACCGGAGAGTGTCTCTGAGCCAG 285
DB 191 GAGTTGCTGGGACAGACGAGGTGAAACGCGGACCGGAGAGTGTCTCTGAGCCAG 250
QY 286 GACAGGTTCTAAGGTCCTGACGCGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGTAC 345
DB 251 GACAGGTTCTAAGGTCCTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTAC 310
QY 346 AATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACTCTGAAGACATC 405
DB 311 AATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACTCTGAAGACATC 370
QY 406 GTGAGGAGCAAAACGCTGAGAGTGCAGCTTATGATTTGTGACACACTCAAGTTACCA 465
DB 371 GTGAGGAGCAAAACGCTGAGAGTGCAGCTTATGATTTGTGACACACTCAAGTTACCA 430
QY 466 GAGACCAAGGTGCTTACCGTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCTAC 525
DB 431 GAGACCAAGGTGCTTACCGTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCTAC 490
QY 526 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGTGAGACACGACTCCGACGTC 585
DB 491 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGTGAGACACGACTCCGACGTC 550
QY 586 AGGCTGTCTGAAGAGTTCTCCGAGCGTGGCGCCGAGGAGGAGGAGGAGGAGGAGTCTG 645
DB 551 AGGCTGTCTGAAGAGTTCTCCGAGCGTGGCGCCGAGGAGGAGGAGGAGGAGGAGTCTG 610
QY 646 AGCCAGTACACCACTTGTGAGAGCGGCGCTTGTGAGAGTCTGCTGCCGACAAAGAG 705
DB 611 AGCCAGTACACCACTTGTGAGAGCGGCGCTTGTGAGAGTCTGCTGCCGACAAAGAG 670
QY 706 TATGCCGATGATCATCCAGAGAGTGAACAATATGTTGTCATCAACTGATCTGTG 765
DB 671 TATGCCGATGATCATCCAGAGAGTGAACAATATGTTGTCATCAACTGATCTGTG 730
QY 766 GAGCAATCCAGAGATCTGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 825
DB 731 GAGCAATCCAGAGATCTGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 790
QY 826 GGGCGGAGTACAAACGACCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGGAGTGTGACCT 885
DB 791 GGGCGGAGTACAAACGACCTTTTCTGAGCCAGGAGGAGGAGGAGGAGGAGTGTGACCT 850
QY 886 GGCAGAGGTCATCATTTGAGTGCAGAGCAGACCCACTGAGGAGGAGGAGGAGGAGGAGT 945
DB 851 GGCAGAGGTCATCATTTGAGTGCAGAGCAGACCCACTGAGGAGGAGGAGGAGGAGGAGT 910
QY 946 GGCAGAGTCTCCGCGCGGATGTGTTTCAAGGAGTGAAGCTTGGAGAGCCGACCCACAC 1005
DB 911 GGCAGAGTCTCCGCGCGGATGTGTTTCAAGGAGTGAAGCTTGGAGAGCCGACCCACAC 970

QY 1006 CCACTGCTTCTCTCGGCGGACCCAGGGAGTGTAGACGAGGCGTTCTTCACTCAG 1065
DB 971 CCACTGCTTCTCTCGGCGGACCCAGGGAGTGTAGACGAGGCGTTCTTCACTCAG 1030
QY 1066 GAGTGAAGAACTCAGATGTGTCTCAGACTCACTTCTGCGGACACAGAGCGTTCT 1125
DB 1031 GAGTGAAGAACTCAGATGTGTCTCAGACTCACTTCTGCGGACACAGAGCGTTCT 1090
QY 1126 GAGTGTTCAGCCACTTATGCTGTTTAAAGATCCCTCTAGGTCACTGAAAT 1185
DB 1091 GAGTGTTCAGCCACTTATGCTGTTTAAAGATCCCTCTAGGTCACTGAAAT 1150
QY 1186 GCCCAGAAATGTGAGGAGCGTGGAGGCTTCTGTGAGGAATGTAGGACATTTATGG 1245
DB 1151 GCCCAGAAATGTGAGGAGCGTGGAGGCTTCTGTGAGGAATGTAGGACATTTATGG 1210
QY 1246 GAAATTTGAGGAGACAGCTTACAGACTGCTGCGCTGATGTTTGTGACAGTGAACCA 1305
DB 1211 GAAATTTGAGGAGACAGCTTACAGACTGCTGCGCTGATGTTTGTGACAGTGAACCA 1270
QY 1306 CAGTGGAGAGAGTGTTCAGTCTGATCTGTTTTCACACTCACTCAACATTAATCA 1365
DB 1271 CAGTGGAGAGAGTGTTCAGTCTGATCTGTTTTCACACTCACTCAACATTAATCA 1330
QY 1366 AAAGTTTGTGAGAGTGTTCAGTCTGATCTGTTTTCAGTCTGATCTGTTTTCAGT 1425
DB 1331 AAAGTTTGTGAGAGTGTTCAGTCTGATCTGTTTTCAGTCTGATCTGTTTTCAGT 1380
QY 1426 TTCTGTTTCAATACAGAGGCTGTTGCTTCAAACTTATTTGATGACCCAGTGGT 1485
DB 1391 TTCTGTTTCAATACAGAGGCTGTTGCTTCAAACTTATTTATGACCCAGTGGT 1450
QY 1486 TTGAGTTCAGCGCTGCTTACAGCGATATGAGGAGCCACTGAGGAGTGTTCCTCC 1545
DB 1451 TTGAGTTCAGCGCTGCTTACAGCGATATGAGGAGCCACTGAGGAGTGTTCCTCC 1510
QY 1546 GCTTGTGCTTAAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1605
DB 1511 GCTTGTGCTTAAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1570
QY 1606 CAGGCTTGTGCGGGGCCAG 1624
DB 1571 CAGGCTTGTGCGGGGCCAG 1589

RESULT 3
AAH15621
ID AAH15621 standard; cDNA; 2160 BP.
XX
AC AAH15621;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13951.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-018776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX

PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 13951; 2537bp + CD ROM; English.
XX
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

Query Match 80.3%; Score 1304; DB 22; Length 2160;
Best Local Similarity 100.0%; Pred No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGGAAAGGGGCGGCGCGGAGCCGATGCGCGAGCGGAGCGGAGCGGAGTGGCTTGGCG 105
DB 47 GGGGAAAGGGGCGGCGCGGAGCCGATGCGCGAGCGGAGCGGAGCGGAGTGGCTTGGCG 106
QY 106 GAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGCGGCGCTTC 165
DB 107 GAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGCGGCGCTTC 166
QY 166 CTGATAGGGGTGAGCGGCGGCACTGCAGCGGAGGAGTGCACGCTGTGTGAGGAAGATCATG 225
DB 167 CTGATAGGGGTGAGCGGCGGCACTGCAGCGGAGGAGTGCACGCTGTGTGAGGAAGATCATG 226
QY 226 GAGTGTGCGGAGCAGAGCGAGTGTGAACAGCGGAGCGGAGGCTGTCTGAGCCAG 285
DB 227 GAGTGTGCGGAGCAGAGCGAGTGTGAACAGCGGAGCGGAGGCTGTCTGAGCCAG 286
QY 286 GACAGGTTTCAAGGTCTCTGAGCGCAGCAGCAAGAGGCCCTTGAAGAAGACGTAC 345
DB 287 GACAGGTTTCAAGGTCTCTGAGCGCAGCAGCAAGAGGCCCTTGAAGAAGACGTAC 346
QY 346 AATTTTGACATGCAATGCTTGTATGATTTATGATGACAGCGACTGGAAGACATC 405
DB 347 AATTTTGACATGCAATGCTTGTATGATTTATGATGACAGCGACTGGAAGACATC 406
QY 406 GTGAGGCGCAAAAGCGTGTGAGTGTGCGACCTATGATTTGTGACACTCAAGTTACCA 465
DB 407 GTGAGGCGCAAAAGCGTGTGAGTGTGCGACCTATGATTTGTGACACTCAAGTTACCA 466
QY 466 GAGACCAAGGTGTCTTACCTCTGCGAGCGTGTCTGTGTTGAGGCGCATCTTGTTCTAC 525
DB 467 GAGACCAAGGTGTCTTACCTCTGCGAGCGTGTCTGTGTTGAGGCGCATCTTGTTCTAC 526

Dh 287 GACAGTTCCTACCAAGTCTCTGACGGCAGAGCAGAGGCCAAGCCCTGAAAAGACAGTAC 346
Qy 346 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGGACTCTGAGAATC 405
Db 347 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGGACTCTGAGAATC 406
Qy 406 GTGAGGGCAAAACGGTGAAGTCCGACCTATGATTTTGTGACACTCAAGTTACCA 465
Db 407 GTGAGGGCAAAACGGTGAAGTCCGACCTATGATTTTGTGACACTCAAGTTACCA 466
Qy 466 GAGACCAAGGTGTGTACCTTCGGACGTGGTCTGTTGAGGGATCTTGTGTCTAC 525
Db 467 GAGACCAAGGTGTGTACCTTCGGACGTGGTCTGTTGAGGGATCTTGTGTCTAC 526
Qy 526 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGAGACCCGACCTCCGAGTC 585
Db 527 AGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGAGACCCGACCTCCGAGTC 586
Qy 586 AGGCTGTCTGAGAAGTCTCCGGGACGTGGCCGAGGGAGGACCTTGACAGATTCTG 645
Db 587 AGGCTGTCTGAGAAGTCTCCGGGACGTGGCCGAGGGAGGACCTTGACAGATTCTG 646
Qy 646 AGCAGTACACCATCTTGTGAAGCCGCTTCGAGGAGTTCTGCTCCGACAAAG 705
Db 647 AGCAGTACACCATCTTGTGAAGCCGCTTCGAGGAGTTCTGCTCCGACAAAG 706
Qy 706 TATGCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATCACTGATCTG 765
Db 707 TATGCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATCACTGATCTG 766
Qy 766 CAGCATCCAGAGCATTTCTGATGTGACATCTTGCAATGGACCGAGAGGGTCCAT 825
Db 767 CAGCATCCAGAGCATTTCTGATGTGACATCTTGCAATGGACCGAGAGGGTCCAT 826
Qy 826 GGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGGGACCACTCTGGATCTGACCTCT 885
Db 827 GGGGGAGCTCAAGCGGACCTTTTCTGAGCAGGGGACCACTCTGGATCTGACCTCT 886
Qy 886 GGCAAAACGGTCACTTTGATGTCCAGCAGACAGCCCACTAGGGGCTCGAGAGCTCAG 945
Db 887 GGCAAAACGGTCACTTTGATGTCCAGCAGACAGCCCACTAGGGGCTCGAGAGCTCAG 946
Qy 946 GGCAAGTCTCCCGCGGATGTGTTCAGG 977
Db 947 GGCAAGTCTCCCGCGGATGTGTTCAGG 978

RESULT 5
AAL36439
ID AAL36439 standard; DNA; 9732 BP.
XX
AC AAL36439;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2804.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249284.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251866.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451937/48.
 XX
 PT Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -

XX
 PS Example 2; SEQ ID NO 2804; 781bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAJ34669-AAJ37666) and proteins
 CC (ABR3087-ABR4109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
 Query Match 54.2%; Score 880; DB 22; Length 9732;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 745 GTTGCCATCAACCTGATGTCAGCAGCATCCAGAGCATTCGATGATGATCTGCAAA 804
 DB 3901 GTTGCCATCAACCTGATGTCAGCAGCATCCAGAGCATTCGATGATGATCTGCAAA 3960
 QY 805 TGGCACCAGAGAGAGGTCCATGAGGCGAGCTCAACAGGACCTTTCTGAGCCAGGGGAGC 864
 DB 3961 TGGCACCAGAGAGAGGTCCATGAGGCGAGCTCAACAGGACCTTTCTGAGCCAGGGGAGC 4020
 QY 865 CACCTTGGATGTCATCTCTGCGCAACCGTCATCATTTGATGTCAGCAGCAGACCCGAC 924
 DB 4021 CACCTTGGATGTCATCTCTGCGCAACCGTCATCATTTGATGTCAGCAGCAGACCCGAC 4080
 QY 925 TGAGGGGCTGCGAGGCTCAGGGCAGGCTCCCGCCGCGATGTGTGTCAGGGATCGAG 984
 DB 4081 TGAGGGGCTGCGAGGCTCAGGGCAGGCTCCCGCCGCGATGTGTGTCAGGGATCGAG 4140
 QY 985 CTTGGGAGCGCCACCCACCACTGCTCTCTGCGGCAACCCAGGGGAGGTGTAGC 1044
 DB 4141 CTTGGGAGCGCCACCCACCACTGCTCTCTGCGGCAACCCAGGGGAGGTGTAGC 4200
 QY 1045 AGCGAGGCTTCTCTCACTAGAGTGAAGTGAATGTCATCTGACTCACTTGTCT 1104
 DB 4201 AGCGAGGCTTCTCTCACTAGAGTGAAGTGAATGTCATCTGACTCACTTGTCT 4260
 QY 1105 GGGACACTGACAGGGGCTCTGAGGTTTCAGCACTTGTGCTGTTGGGTTTAAGAT 1164
 DB 4261 GGGACACTGACAGGGGCTCTGAGGTTTCAGCACTTGTGCTGTTGGGTTTAAGAT 4320
 QY 1165 CCCTTAGGTCACTGAGAAATGCCAGAAATGTCAGAGGCTGAGAGCTTCTGTGAG 1224
 DB 4321 CCCTTAGGTCACTGAGAAATGCCAGAAATGTCAGAGGCTGAGAGCTTCTGTGAG 4380
 QY 1225 GAATGTGAGGCACATTATTTGGGAAATTTAGAGAGACGCTTGAACACTGGCTGTGAT 1284
 DB 4381 GAATGTGAGGCACATTATTTGGGAAATTTAGAGAGACGCTTGAACACTGGCTGTGAT 4440
 QY 1285 GTTTTGTGACAGTGAACCCACAGTGGGAGAGGTTTTCAGTGTGATGTGTTCTTA 1344
 DB 4441 GTTTTGTGACAGTGAACCCACAGTGGGAGAGGTTTTCAGTGTGATGTGTTCTTA 4500
 QY 1345 CACACTCAGACATTAAGTAAAGTTTGTGAACAGTACTTCTTTTACATGTTA 1404
 DB 4501 CACACTCAGACATTAAGTAAAGTTTGTGAACAGTACTTCTTTTACATGTTA 4560

QY 1405 CATGTCCTCATGATTTTCTGTTTCTGTTTCATTAACACAAAGCTGGTTGGCCTTAACAAC 1464
 DB 4561 CATGTCCTCATGATTTTCTGTTTCTGTTTCTAATTAACAAAGCTGGTTGTGCTTAACAAC 4620
 QY 1465 CTAATTTATATACCCAGTGTGTTTGAGTTCACAGCGTGGCTTAACGATATGGGAGCCAC 1544
 DB 4621 CTAATTTATATACCCAGTGTGTTTGAGTTCACAGCGTGGCTTAACGATATGGGAGCCAC 4680
 QY 1525 TGAGGAGATGTTTCCCGCTTGCTGTCGCTTAAGAGGAGAGGAGGAGGAGTGCCT 1584
 DB 4681 TGAGGAGATGTTTCCCGCTTGCTGTCGCTTAAGAGGAGAGGAGGAGGAGTGCCT 4740
 QY 1585 GGAAGCACCCAGCATACACACCAGGCTTGTGCGGGCCAG 1624
 DB 4741 GGAAGCACCCAGCATACACACCAGGCTTGTGCGGGCCAG 4780
 RESULT 6
 ABX59427
 ID ABX59427 standard; cDNA; 9732 BP.
 AC ABX59427;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE cDNA encoding novel human musculoskeletal system antigen #1771.
 XX
 KM Gene; 68; musculoskeletal system antigen; cancer; metastasis;
 KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KM post-operative tissue repair; limb regeneration; neuronal growth;
 KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KM AIDS-related complex; chondrocyte growth; bone regeneration;
 KM periodontal regeneration; tissue transport; bone graft; skin aging;
 KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KM cell growth; organ transplant; cell differentiation; body height;
 KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KM pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
 KM depression; tendency for violence; pain; reproductive capability;
 KM hormone level; endocrine level; appetite; libido; memory; stress;
 KM storage capability; fat content; lipid content; protein content;
 KM carbohydrate content; vitamin content; cofactor content;
 KM nutritional component.
 XX
 OS Homo sapiens.
 XX
 PN US2002147140-A1.
 PD 10-OCT-2002.
 XX
 PF 17-JAN-2001; 2001US-0764877.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 17-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.

PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239935P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-128199/12.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 PT associated polypeptides, useful for detecting disorders, e.g. cancer -
 PS
 PS Disclosure; SEQ ID NO 2804; 321pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid; stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, hematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's mental state or physical state by influencing biorhythms;
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human

CC musculoskeletal system antigen.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.

XX Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Query March 54.2%; Score 880; DB 25; Length 9732;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTCCATCAACCTGATGTCAGACATCCAGGACATTTGATGATGATCGACAA 804

Db 3901 GTTCCATCAACCTGATGTCAGACATCCAGGACATTTGATGATGATCGACAA 3960

QY 805 TGGCACCAGAGAGGGTCCATGAGCGGAGCTACAAAGCGACCTTTCTAGCCAGGGGAC 864

Db 3961 TGGCACCAGAGAGGGTCCATGAGCGGAGCTACAAAGCGACCTTTCTAGCCAGGGGAC 4020

QY 865 CACCTTGAGATGTCGACCTCTGCGCAACGGTCATTTGAGTCAGACGACACCCAC 924

Db 4021 CACCTTGAGATGTCGACCTCTGCGCAACGGTCATTTGAGTCAGACGACACCCAC 4080

QY 925 TGAAGGAGCTGCGGAGCTCAGAGGAGGTCTCCGCGGAGTGTGTTGAGGAGCTGAG 984

Db 4081 TGAAGGAGCTGCGGAGCTCAGAGGAGGTCTCCGCGGAGTGTGTTGAGGAGCTGAG 4140

QY 985 CTTGGGAGCGGCCAACCCACACCACTGCTCTCTGCGGCGGACCCAGGGGAGTGTAGC 1044

Db 4141 CTTGGGAGCGGCCAACCCACACCACTGCTCTCTGCGGCGGACCCAGGGGAGTGTAGC 4200

QY 1045 AGGAGGCGCTTCTCTCACTCAGAGATGAGAACTCAGATGTCATCTCAACTTCT 1104

Db 4201 AGGAGGCGCTTCTCTCACTCAGAGATGAGAACTCAGATGTCATCTCAACTTCT 4260

QY 1105 GGGACATGACAGGCGTTCTGAGGTTTTCAGCACTTGAAGCTGTTGCGGTTTAAAGAT 1164

Db 4261 GGGACATGACAGGCGTTCTGAGGTTTTCAGCACTTGAAGCTGTTGCGGTTTAAAGAT 4320

QY 1165 CCCTCTAGGTCACAGAAATGCCACAGATGTCAGAGAGCCCTGGAGGCTTCTGTAG 1224

Db 4321 CCCTCTAGGTCACAGAAATGCCACAGATGTCAGAGAGCCCTGGAGGCTTCTGTAG 4380

QY 1225 GAATGTAGGACATTTATTTGGGGAAATTGAGGACACGCTTACACCTGCTGCTGAT 1284

Db 4381 GAATGTAGGACATTTATTTGGGGAAATTGAGGACACGCTTACACCTGCTGCTGAT 4440

QY 1285 GTTTGTTGACAGGAAACCCACAGTGGGAGAGATTTTTCAGTCTGATCTGTTCTTA 1344

Db 4441 GTTTGTTGACAGGAAACCCACAGTGGGAGAGATTTTTCAGTCTGATCTGTTCTTA 4500

QY 1345 CACACTCACAACATTAACCTCAAAAGTTTGTGAACAAATCTTTCTTTTATCATGTTA 1404

Db 4501 CACACTCACAACATTAACCTCAAAAGTTTGTGAACAAATCTTTCTTTTATCATGTTA 4560

QY 1405 CAGTCTCTCATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGCTCTCAAAAC 1464

Db 4561 CAGTCTCTCATGTTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGTGCTCTCAAAAC 4620

QY 1465 CTAAATTCATGACCCAGTGTGTTGTCAGTCCAGCGTGGCTTACAGGATATGGGGAGCAC 1524

Db 4621 CTAAATTCATGACCCAGTGTGTTGTCAGTCCAGCGTGGCTTACAGGATATGGGGAGCAC 4680

QY 1525 TGAAGGATGTTTCTCCCTTGTGCTTAAAGGACAGAAAGCGAGCGGATGCTT 1584

Db 4681 TGAAGGATGTTTCTCCCTTGTGCTTAAAGGACAGAAAGCGAGCGGATGCTT 4740

QY 1585 GGAAGCACCAGATCAACCCAGGCTTGTGCGGGGCGAG 1624

Db 4741 GGAAGCACCAGATCAACCCAGGCTTGTGCGGGGCGAG 4780

RESULT 7

ID AAL36440 standard; DNA; 19125 BP.

XX AAL36440;

DT 08-JAN-2002 (first entry)

DE Human musculoskeletal system related polynucleotide SEQ ID NO 2805.

XX Cyostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;

XX vulnereary; anticoagulant; antibacterial; antifungal; antiparasitic;

XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX neurological disease; infection; human; secreted protein;

XX musculoskeletal system; ds.

XX Homo sapiens.

PN NO200155367-A1.

PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01338.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 05-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX Example 2; SEQ ID NO 2805; 781bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (AAB01087-AAB04109) associated with the musculoskeletal system useful
XX for preventing, treating or ameliorating medical conditions e.g. by
XX protein or gene therapy. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins,
XX antibodies and (ant)agonists are useful in the diagnosis, treatment,
XX and prevention of: (a) cancer, e.g. breast and ovarian cancer and
XX other cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
XX healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
XX and (f) infectious diseases such as viral, bacterial, fungal and
XX parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
SQ
Query Match 54.2%; Score 880; DB 22; Length 19125;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 880; Conservative 0; Indels 0; Gaps 0;
QY 745 GTTGCCATCACTGATGTCGAGCAATCCAGACATTTGATGATGATCTGCATAA 804
DB 5997 GTTGCCATCACTGATGTCGAGCAATCCAGACATTTGATGATGATCTGCATAA 6056
QY 805 TGGCACCAGAGAGGTCCTGCAATGCGCAATGCGGAGTACAAAGCGAATTTTCTGAGCGGAGAC 864
DB 6057 TGGCACCAGAGAGGTCCTGCAATGCGGAGTACAAAGCGAATTTTCTGAGCGGAGAC 6116
QY 865 CACCTGGGATGTCGACCTCTGCAAAAGGTCAATTTGAGTCGAGCAGACCCAC 924
DB 6117 CACCTGGGATGTCGACCTCTGCAAAAGGTCAATTTGAGTCGAGCAGACCCAC 6176
QY 925 TGAAGGGCTGCGAGCTCAGGCGAGGTCTCCCGCGGATGTCGAGGAGCTGAG 984
DB 6177 TGAAGGGCTGCGAGCTCAGGCGAGGTCTCCCGCGGATGTCGAGGAGCTGAG 6236
QY 985 CTTGGGGAGCGCCACCAACCAACCACTGCTTCTCTGCGCGACCCCAAGGGAGTGTAGC 1044

Dh 6237 CTTGGGAGCCGCCACCCACACCTGCTTCTCTCGGCGCACCAGGGAGTGTAGC 6296
Qy 1045 AGGAGGSCCTTCTCACTCAAGAGTGAATCTCAAGTGTCACTGACTCAACTTGGT 1104
Db 6297 AGCGAGSCCTTCTCACTCAAGAGTGAATCTCAAGTGTCACTGACTCAACTTGGT 6356
Qy 1105 GGGACATGACAGGCGCTTCTGAGGTTTTCAGCCACTTAAGCTGCTGGCTTTAAAGAT 1164
Db 6357 GGGACATGACAGGCGCTTCTGAGGTTTTCAGCCACTTAAGCTGCTGGCTTTAAAGAT 6416
Qy 1165 CCCTCTAGTCACTGAGAAATGCCACAAATGTGCAGAAAGCTGGAGGCTTCTGTAG 1224
Db 6417 CCCTCTAGTCACTGAGAAATGCCACAAATGTGCAGAAAGCTGGAGGCTTCTGTAG 6476
Qy 1225 GAATGTAGGACACTTATTGGGAAATTTGAGAGACGCTTAAGCACTGGCTGGCCGAT 1284
Db 6477 GAATGTAGGACACTTATTGGGAAATTTGAGAGACGCTTAAGCACTGGCTGGCCGAT 6536
Qy 1285 GTTTTGTGACAGTGAACCCACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 1344
Db 6537 GTTTTGTGACAGTGAACCCACAGTGGAGAGATTTTTCAGTCTGATCTGGTCTTA 6596
Qy 1345 CACACTCACAACATAAATCTGAAAGTTTGTGAACAAGTCTTCTTTTACATGTTA 1404
Db 6597 CACACTCACAACATAAATCTGAAAGTTTGTGAACAAGTCTTCTTTTACATGTTA 6656
Qy 1405 CAGTCTCAGTGTCTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGGCTTAACAAC 1464
Db 6657 CAGTCTCAGTGTCTTCTGTTTCTGTTTCTATTAACAAGGCTGTGTGGCTTAACAAC 6716
Qy 1465 CTATTTTCATGACCCAGTGTGTTGAGTCCAGCGTGGCTTAACAAGTATGGGAGGCAC 1524
Db 6717 CTATTTTCATGACCCAGTGTGTTGAGTCCAGCGTGGCTTAACAAGTATGGGAGGCAC 6776
Qy 1525 TGAAGGATGTTTTCCTCCCTTGTGCTTGAAGGACAGAGAGCGAGATGCCCT 1584
Db 6777 TGAAGGATGTTTTCCTCCCTTGTGCTTGAAGGACAGAGAGCGAGATGCCCT 6836
Qy 1585 GGAAGCACCCAGATCAACCCAGGCTTGTGCGGGGCGAG 1624
Db 6837 GGAAGCACCCAGATCAACCCAGGCTTGTGCGGGGCGAG 6876

RESULT 8
ABX59428
ID ABX59428 standard, cDNA; 19125 BP.
XX
AC ABX59428;
XX
DT 26-FEB-2003 (first entry)
XX
DE cDNA encoding novel human musculoskeletal system antigen #1772.
XX
XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;
XX periodontal regeneration; tissue transport; bone graft; skin aging;
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
XX cell growth; organ transplant; cell differentiation; body height;
XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
XX pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;
XX depression; tendency for violence; pain; reproductive capability;
XX hormone level; endocrine level; appetite; libido; memory; stress;
XX storage capability; fat content; lipid content; protein content;
XX carbohydrate content; vitamin content; cofactor content;
XX nutritional component.
XX
XX Homo sapiens.
XX
XX US2002147140-A1.

XX 10-OCT-2002.
PD 17-JAN-2001; 2001US-0764877.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234977P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236337P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236820P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239355P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC,
XX WPI, 2003-128199/12.
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
XX associated polypeptides, useful for detecting disorders, e.g. cancer -
XX
XX Disclosure; SEQ ID NO 2805; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
XX sequence encoding musculoskeletal system associated polypeptides useful
XX for detecting disorders, e.g., cancer or cancer metastases, in animals
XX or humans. The nucleic acid: stimulates re-vascularisation of ischaemic

CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, haematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's mental state or physical state by influencing biorhythms,
 CC circadian rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional system components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?docid=20020147140.

XX Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;

CC Query Match 54.2%; Score 880; DB 25; Length 19125;

CC Best Local Similarity 100.0%; Pred. No. 0;

CC Matches 880; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 745 GTTGCAATCAACCGATGTCAGACATCCAGGCAATTGATGGTACATCTGGAAA 804
 DB 5997 GTTGCAATCAACCGATGTCAGACATCCAGGCAATTGATGGTACATCTGGAAA 6056
 QY 805 TGGACCGAGAGGGGTCATAGGCGGAGCTTAAAGCGGACCTTTTGAAGCCAGGGGAC 864
 DB 6057 TGGACCGAGAGGGGTCATAGGCGGAGCTTAAAGCGGACCTTTTGAAGCCAGGGGAC 6116
 QY 865 CACCTTGAGATGTCAGCTCTGGCAAAGGTCAATTGGAGTCAAGCAGCAGACCCAC 924
 DB 6117 CACCTTGAGATGTCAGCTCTGGCAAAGGTCAATTGGAGTCAAGCAGCAGACCCAC 6176
 QY 925 TGAGGGGCTGCGGAGCTCAAGGAGGCTCCCGCCGGGCAATGTCAGGGGAGCTGAG 984
 DB 6177 TGAGGGGCTGCGGAGCTCAAGGAGGCTCCCGCCGGGCAATGTCAGGGGAGCTGAG 6236
 QY 985 CTTGGGGAGCGCCCAACCAACCACTGCTTCTCTCGCGGACCCAGGGGAGGTTAGC 1044
 DB 6237 CTTGGGGAGCGCCCAACCAACCACTGCTTCTCTCGCGGACCCAGGGGAGGTTAGC 6286
 QY 1045 AGCGAGGCTTCTCACTCAGAGAGTGAATCTCAGATGTCACCTCAGACTCAACTGGCT 1104
 DB 6297 AGCGAGGCTTCTCACTCAGAGAGTGAATCTCAGATGTCACCTCAGACTCAACTGGCT 6356
 QY 1105 GGGACACTGACAGAGGCTTCTGAGGTTTTCAGCACTTAGGCTGCTGGGTTAAAGAT 1164
 DB 6357 GGGACACTGACAGAGGCTTCTGAGGTTTTCAGCACTTAGGCTGCTGGGTTAAAGAT 6416
 QY 1165 CCTCTAGGTCACTGAGAAATGCCACAGAAATGTCAGGAAAGCTTGGAGGCTTGTGAG 1224
 DB 6417 CCTCTAGGTCACTGAGAAATGCCACAGAAATGTCAGGAAAGCTTGGAGGCTTGTGAG 6476
 QY 1225 GAATGTGAGGACATTTATTTGGGAAATTTGAGGAGACAGCTTAGACACTGCTGCTGAT 1284
 DB 6477 GAATGTGAGGACATTTATTTGGGAAATTTGAGGAGACAGCTTAGACACTGCTGCTGAT 6536

QY 1285 GTTTGTTGACAGTGAACCCAGAGGAGAGGTTTTTCCAGTCGATCTGTTCTTA 1344
 DB 6537 GTTTGTTGACAGTGAACCCAGAGGAGAGGTTTTTCCAGTCGATCTGTTCTTA 6596
 QY 1345 CACACTCACACATTAACCTAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTA 1404
 DB 6597 CACACTCACACATTAACCTAAAGTTTGTGAACAAGTACTTCTTTTTCATGTTA 6656
 QY 1405 CAGTCTCATGTTTTCTGTTTCTGTTTCATTAACACAGGCTGTTGAGCTTAACAAC 1464
 DB 6657 CAGTCTCATGTTTTCTGTTTCTGTTTTCATTAACAAGGCTGTTGAGCTTAACAAC 6716
 QY 1465 CTAAATTCAGACCCAGTGGTTTTCAGTCAGGAGGCTTACAGGATATGGGAGCCAC 1524
 DB 6717 CTAAATTCAGACCCAGTGGTTTTCAGTCAGGAGGCTTACAGGATATGGGAGCCAC 6776
 QY 1525 TGAGGAGTGTTCCTCCCTTGTGCTTGTAAAGCAGAGAGGAGCGGATGCCCT 1584
 DB 6777 TGAGGAGTGTTCCTCCCTTGTGCTTGTAAAGCAGAGAGGAGCGGATGCCCT 6836
 QY 1585 GGAAGCACCCAGCATACACCCAGGCTTGTGCGGGCCAG 1624
 DB 6837 GGAAGCACCCAGCATACACCCAGGCTTGTGCGGGCCAG 6876

RESULT 9
 AAK98735
 ID AAK98735 standard; DNA; 834 BP.
 AC AAK98735;
 XX
 DT 02-MAY-2002 (first entry)
 XX
 DE DNA of a human uridine kinase (UDK).
 XX
 KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW human colon carcinoma; immunogen; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..780
 FT /tag= a
 FT /partial
 FT /note= "No stop codon"
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI; 2001-626259/72.
 DR P-FSDB; AAO14412.
 XX
 XX Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance -
 XX
 PS Claim 5; Page 22-23; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention

are methods for utilizing uridine kinase polypeptides and polynucleotides in diagnostic assays. The polynucleotides and polypeptides of the invention may be used as diagnostic reagents by detecting mutations in an associated gene. An array of oligonucleotide probes comprising the uridine kinase polynucleotide sequence or fragments thereof can be constructed to conduct efficient screening of genetic mutations, for example. Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. The polynucleotide sequences of the invention can be used for chromosome localisation studies and tissue expression studies. The polypeptides of the invention or fragments thereof may be used as immunogens to produce antibodies. These antibodies may be employed to isolate or identify clones expressing the polypeptide. The polypeptides and polynucleotides of the invention can be used as a vaccine or in gene therapy to treat diseases such as human ovarian cancer, human colon carcinomas, and immunological disorders. This polynucleotide sequence represents the DNA of a human uridine kinase of the invention.

Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 48.2%; Score 783; DB 23; Length 834;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 94 ATGGCTTGGGGGAGGCGAAGCTGAGAGCCCGCGCGGAGGCGGACCTCCGAC 153
DB 1 ATGGCTTGGGGGAGGCGAAGCTGAGAGCCCGCGCGGAGGCGGACCTCCGAC 60
QY 154 CAGCGGCTTCTGATAGGGGTGAGCGGCGCACTCCAGCGGGAAGTGCAGCGTGT 213
DB 61 CAGCGGCTTCTGATAGGGGTGAGCGGCGCACTCCAGCGGGAAGTGCAGCGTGT 120
QY 214 GAGAGATCATGAGATTCTGGGACAGACGAGGTGGAACGCGGCGGGAAGTGTGTC 273
DB 121 GAGAGATCATGAGATTCTGGGACAGACGAGGTGGAACGCGGCGGGAAGTGTGTC 180
QY 274 ATCTGAGCGGAGGAGGTTCTCAAGGTCCTGACGCGGAGGAGGAGGAGGCTTGG 333
DB 181 ATCTGAGCGGAGGAGGTTCTCAAGGTCCTGACGCGGAGGAGGAGGAGGCTTGG 240
QY 334 AAAGGAGTACAAATTTTGAACATCCAGATGCTTTGATGATGATGATGATGATGATG 393
DB 241 AAAGGAGTACAAATTTTGAACATCCAGATGCTTTGATGATGATGATGATGATGATG 300
QY 394 CTGAGAGATCTGTGAGGGCAAAACGTTGAGGTGCGAATTATGATTTTGTGACACAC 453
DB 301 CTGAGAGATCTGTGAGGGCAAAACGTTGAGGTGCGAATTATGATTTTGTGACACAC 360
QY 454 TCAAGTTTACCAAGACACAGGTGTCTACCTTCGCGAGCTGTCTGTTTGAAGGATC 513
DB 361 TCAAGTTTACCAAGACACAGGTGTCTACCTTCGCGAGCTGTCTGTTTGAAGGATC 420
QY 514 TTGGTGTCTACAGCCGAGGATCCGAGCATGTTCCAGCTGGCGCTCTTCGTTGAGCAC 573
DB 421 TTGGTGTCTACAGCCGAGGATCCGAGCATGTTCCAGCTGGCGCTCTTCGTTGAGCAC 480
QY 574 GACTCCGACGTGAGCTGTCTGAAAGATTCTCCGAGACGTGCGCCGAGGAGGAGCTG 633
DB 481 GACTCCGACGTGAGCTGTCTGAAAGATTCTCCGAGACGTGCGCCGAGGAGGAGCTG 540
QY 634 GAGCAATTTTGAAGCAGTACACCACTTCTGTGAAGCGGCTTTCAGAGGATTTGCTG 693
DB 541 GAGCAATTTTGAAGCAGTACACCACTTCTGTGAAGCGGCTTTCAGAGGATTTGCTG 600
QY 694 CCAGCAAAAGATATGCGGATGATCATCCAGAGAGTGAAGCAATATGTTGCTGATC 753
DB 601 CCAGCAAAAGATATGCGGATGATCATCCAGAGAGTGAAGCAATATGTTGCTGATC 660
QY 754 AACCTGATGTCGAGCATCCAGAGCATTTCTGAATGTGATCATCTTGAATGAGACCGA 813
DB 661 AACCTGATGTCGAGCATCCAGAGCATTTCTGAATGTGATCATCTTGAATGAGACCGA 720

```

```

QY 814 GGAGGTCCAAATGGGCGAGCTACAGCGACCTTTTGTGAGCGAGGAGGACCTGGG 873
DB 721 GGAGGTCCAAATGGGCGAGCTACAGCGACCTTTTGTGAGCGAGGAGGACCTGGG 780
QY 874 ATGCTGACCTTGGCAACGCTGACATTTGAGTCCAGAGCAGACCCCACTGA 927
DB 781 ATGCTGACCTTGGCAACGCTGACATTTGAGTCCAGAGCAGACCCCACTGA 834

```

RESULT 10

ABL89762 standard; cDNA; 1396 BP.

ABL89762;

24-MAY-2002 (first entry)

Human polynucleotide SEQ ID NO 324.

Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer; vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein; gene; ss.

Homo sapiens.

WO200190304-A2.

29-NOV-2001.

18-MAY-2001; 2001MO-US16450.

19-MAY-2000; 2000US-205515P.

(HUMA-) HUMAN GENOME SCI INC.

Birse CE, Rosen CA;

WPI; 2002-122018/16.

P-FSDB; ABB89353.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -

Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABL9085) and proteins (ABB89040-ABB9044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;

Query Match 38.7%; Score 628; DB 24; Length 1396;

Best Local Similarity 99.3%; Pred. No. 8; 4e-307;

Matches 1288; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

```
QY 54 GGCGGCGCGGGGACCCGATGCGCGGAGCGAGGCGGAGATGCTTGGCGGGAGGGA 113
Db 1 GGCGGCGCGGGGACCCGATGCGCGGAGCGAGGCGGAGATGCTTGGCGGGAGGGA 60
QY 114 AGACTGCGAGAGCCCGCGCGGAGCCGACGCTCCGACCAAGCGGCCCTTCTGATAGG 173
Db 61 AGACTGCGAGAGCCCGCGCGGAGCCGACGCTCCGACCAAGCGGCCCTTCTGATAGG 120
QY 174 GGTGAGCGCGGGGACTGCGACGCGGGAAGTCAGCCGTGTGTAGAAATGATGAGATGCT 233
Db 121 GGTGAGCGCGGGGACTGCGACGCGGGAAGTCAGCCGTGTGTAGAAATGATGAGATGCT 180
QY 234 GGGACAGAACGAGGTGGAACAGCGGCAAGGAGGTGCTCACTGAGCCAGACAGATT 293
Db 181 GGGACAGAACGAGGTGGAACAGCGGCAAGGAGGTGCTCACTGAGCCAGACAGATT 240
QY 294 CTACAGAGTCTGTACGCGGAGAGCAGAGGCCAAGGCTTGAAGAGACATATTTTGA 353
Db 241 CTACAGAGTCTGTACGCGGAGAGCAGAGGCCAAGGCTTGAAGAGACATATTTTGA 300
QY 354 CCATCCAGATGCTTGTATGATTTGATGACAGAGCTGTGAAGAACATCGTGGAGGG 413
Db 301 CCATCCAGATGCTTGTATGATTTGATGACAGAGCTGTGAAGAACATCGTGGAGGG 360
QY 414 CAAACGCTGAGGAGTGCAGCCTATGATTTTGTGACACACTCAAGTTTACAGAGACAC 473
Db 361 CAAACGCTGAGGAGTGCAGCCTATGATTTTGTGACACACTCAAGTTTACAGAGACAC 420
QY 474 GGTGTCTACCTCTGCGGAGCTGTGTCTGTTGAGGGACCTTGTGTCTTACACCGAGA 533
Db 421 GGTGTCTACCTCTGCGGAGCTGTGTCTGTTGAGGGACCTTGTGTCTTACACCGAGA 480
QY 534 GATCGGGGACATGTTTCCAGCTGGGCTCTTGTGTGACACCGACCTCCAGCGTGTGC 593
Db 481 GATCGGGGACATGTTTCCAGCTGGGCTCTTGTGTGACACCGACCTCCAGCGTGTGC 540
QY 594 TCGAAGAGTCTTCCGCGGAGCTGCGCCGAGGAGGAGGAGCTGAGAGATTTGACGACATA 653
Db 541 TCGAAGAGTCTTCCGCGGAGCTGCGCCGAGGAGGAGGAGCTGAGAGATTTGACGACATA 599
QY 654 CACCACTTGTGTGAAGCGGCTTTCGAGAGATTTGCTGCTGCCGACAAAGATATGCCGA 713
Db 600 CACCACTTGTGTGAAGCGGCTTTCGAGAGATTTGCTGCTGCCGACAAAGATATGCCGA 659
QY 714 TGTGATATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGTTCGCGACGACAT 773
Db 660 TGTGATATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGTTCGCGACGACAT 719
QY 774 CCAGAGCATTTCTGAATGATGATCTGCAATATGACACCGAGAGGATCAATGGGCGAG 833
Db 720 CCAGAGCATTTCTGAATGATGATCTGCAATATGACACCGAGAGGATCAATGGGCGAG 779
QY 834 CTACAGAGGACCTTTTCTGAGCCAGGAGACACCTTGGATGCTGACTTGGCAACG 893
Db 780 -TACAAAGCGGACCTTTTCTTAGCCAGGAGACACCTTGGATGCTGACTTGGCAACG 838
QY 894 GTCAATTTTGGAGTCCAGAGACAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGTGC 953
Db 839 GTCAATTTTGGAGTCCAGAGACAGACCCCACTGAGGGGCTGCGAGGCTCAAGGGCAGTGC 898
QY 954 TCCGCGCGGAGATGTGTTCAGGAGACTGA-GCCTGGGGAAGCCCAACCAACCACTGCG 1012
Db 899 TCCGCGCGGAGATGTGTTCAGGAGATTKAGGCTTGGGGAAGCCCAACCAACCACTGCG 958
QY 1013 TTCTCTTCGGGCAACCCAGGGGAGTGTAGACAGAGGCTTCTCACTCAGAGAGTGA 1072
Db 959 TTCTCTTCGGGCAACCCAGGGGAGTGTAGACAGAGGCTTCTCACTCAGAGAGTGA 1018
QY 1073 AACTCAGATGTGACTCAGACTCAACTGTGTGGGACACTGACAGGGGTTTCTGAGGTTT 1132
Db 1019 AACTCAGATGTGACTCAGACTCAACTGTGTGGGACACTGACAGGGGTTTCTGAGGTTT 1078
```

```
QY 1133 TCAGCCACTTAGGCTGCTGTGGCTTTAAAGATCCCTTAGTCACTGAGAAATGCCAG 1192
Db 1079 TCAGCCACTTAGGCTGCTGTGGCTTTAAAGATCCCTTAGTCACTGAGAAATGCCAG 1138
QY 1193 AATGTGAGAGACCTTGGAGGCTTCTGTGAGAAATGTGAGGCAATTAATGGGAAATT 1252
Db 1139 AATGTGAGAGACCTTGGAGGCTTCTGTGAGAAATGTGAGGCAATTAATGGGAAATT 1198
QY 1253 GAGGAGACAGCCTTAGACACTGAGGCTGAGATGTTTGTGACAGTGAACCCACAGTGG 1312
Db 1199 GAGGAGACAGCCTTAGACACTGAGGCTGAGATGTTTGTGACAGTGAACCCACAGTGG 1258
QY 1313 AGAGAGTTTTCAGTCTGATCTGTGTTCTTACACAC 1349
Db 1259 AGAGAGTTTTCAGTCTGATCTGTGTTCTTACACAC 1295

RESULT 11
AAH04832
ID AAH04832 standard; cDNA; 753 BP.
XX
AC AAH04832;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:1667.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JUN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000UP-0183767.
PR 09-JUN-2000; 2000UP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
```


QY 865 CACCCTGGAGTGTGACCTCTGGCAACGATCATTTGGATGTCAGACGAGACCCAC 924
 |||||
 DB 776 CACCTGGAGTGTGACCTCTGGCAACGATCATTTGGATGTCAGACGAGACCCAC 835
 |||||
 QY 925 TGAAGGAGTGTGACCTCTGGCAACGATCATTTGGATGTCAGACGAGACCCAC 984
 |||||
 DB 836 TGAGGAGTGTGACCTCTGGCAACGATCATTTGGATGTCAGACGAGACCCAC 895
 |||||
 QY 985 CCTGG 989
 |||||
 DB 896 CCTGG 900
 |||||

RESULT 13
 ABS51801
 ID ABS51801 standard; cDNA, 734 BP.
 XX
 AC ABS51801;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human mdt cDNA Incyte ID No: LI:235557.12:2001JAN12.
 XX
 KW Human; molecule for disease detection and treatment; MDDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytosolic; antiatherosclerotic; anti-inflammatory;
 KW hepatotropic; immunosuppressive; antiasthmatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W020255738-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US01008.
 XX
 PR 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCYTE) INCYTE GENOMICS INC.
 XX
 PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL,
 PI Dam TC, Liu TP, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ,
 PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA,
 XX
 DR WPI: 2002-590679/63.
 DR P-PSDB; ABG70328.
 XX
 OS New disease detection and treatment molecule (MDDT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDT expression, e.g. autoimmune or
 PT inflammatory disorders
 XX
 PS Claim 1; Page 98; 129p; English.
 XX
 CC The present invention relates to the isolation of novel human
 CC molecules for disease detection and treatment (MDDT), and the
 CC polynucleotide sequences (mdt) encoding them. The MDDT polypeptides
 CC may be used to screen for molecules that bind to, or are bound by the
 CC encoded polypeptides, and to develop a transcript image of a tissue or
 CC cell type. Probes comprising at least 20 nucleotides of the mdt
 CC polynucleotide may be used to assess the toxicity of a test compound.
 CC The MDDT polypeptides and mdt polynucleotides are useful in the
 CC diagnosis, study, prevention and treatment of diseases associated with
 CC the expression of molecules for disease detection and treatment. Such
 CC disorders include cell proliferative disorders (e.g. arteriosclerosis,

CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
 CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdt
 CC polynucleotides may also be used as molecule markers, in microarrays,
 CC and in somatic or germline gene therapy. ABS51779-ABS51814 encode
 CC the MDDT proteins of the invention.
 XX
 SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;

Query Match 20.8%; Score 338; DB 24; Length 734;
 Best Local Similarity 100.0%; Pred. No. 3.2e-160;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 CAGATGCTTTGATTAATGATTTGATGACACGACTCTGAAGACATCTGTGAGGCAAAA 418
 |||||
 DB 12 CAGATGCTTTGATTAATGATTTGATGACACGACTCTGAAGACATCTGTGAGGCAAAA 71
 |||||
 QY 419 CGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCGGTGG 478
 |||||
 DB 72 CGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCGGTGG 131
 |||||
 QY 479 TCTACCTGCGGAGCGGTCTCTTTGAGGAGCACTTTGTGTTCTACGACGAGATCC 538
 |||||
 DB 132 TCTACCTGCGGAGCGGTCTCTTTGAGGAGCACTTTGTGTTCTACGACGAGATCC 191
 |||||
 QY 539 GGGACATGTTCCACCTGCGGCTCTTGTGAGACACGACTCCGACGCTGTCTGAA 598
 |||||
 DB 192 GGGACATGTTCCACCTGCGGCTCTTGTGAGACACGACTCCGACGCTGTCTGAA 251
 |||||
 QY 599 GAGTTCTCCGGAGCGGTCTCTTTGAGGAGCACTTTGTGAGACGAGATTTGACGACATACCA 658
 |||||
 DB 252 GAGTTCTCCGGAGCGGTCTCTTTGAGGAGCACTTTGTGAGACGAGATTTGACGACATACCA 311
 |||||
 QY 659 CTTTGGAGAGCGGCTCTTGTGAGGAGCACTTTGTGAGACGAGATTTGACGACATACCA 696
 |||||
 DB 312 CTTTGGAGAGCGGCTCTTGTGAGGAGCACTTTGTGAGACGAGATTTGACGACATACCA 349
 |||||

RESULT 14
 AAL35057
 ID AAL35057 standard; cDNA; 222 BP.
 XX
 AC AAL35057;
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Human musculoskeletal system related polynucleotide SEQ ID NO 399.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein;
 KW musculoskeletal system; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200155367-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01338.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216860.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 16-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226861.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-02289924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM,
 PI WPI; 2001-451937/48.
 XX P-PSDB; ABB03475.
 DR
 DR
 XX
 PT Isolated polypeptide for treating, preventing and/ or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 XX
 XX
 PS Claim 1; SEQ ID NO 399; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAU34669-AAU37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment

CC	and prevention of:	(a) cancer, e.g. breast and ovarian cancer and
CC	other cancers of the adrenal gland, bone, bone marrow, breast,	
CC	gastrointestinal tract, liver, lung, or urogenital;	(b) immune
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
CC	anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,	
CC	multiple sclerosis, rheumatoid arthritis and ulcerative colitis;	
CC	(c) cardiovascular disorders such as myocardial ischaemias; (d) wound	
CC	healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;	
CC	and (f) infectious diseases such as viral, bacterial, fungal and	
CC	parasitic infections.	
CC	Note: The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 222 BP; 44 A; 63 C; 75 G; 38 T; 2 other;	
OY		
DG	1521 CCACGAGGAGGTTCCTCCCCCTGCCTTGCGCTTAAGCAGAGGCGGCGA 1578 32 CCACGAGGAGGTTCCTCCCCCTGCCTTGCGCTTAAGCAGAGGCGGCGA 89	
RESULT 15		
ID	ABX58045	
XX	ABX58045 standard; cDNA; 222 BP.	
AC	ABX58045;	
XX		
DT	26-FEB-2003 (first entry)	
XX		
DE	cDNA encoding novel human musculoskeletal system antigen #369.	
XX		
KM	Gene; ss; musculoskeletal system antigen; cancer; metastasis;	
KM	re-vascularisation; thrombosis; arteriosclerosis; mineral content;	
KM	cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;	
KM	post-operative tissue repair; limb regeneration; neuronal growth;	
KM	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;	
KM	AIDS-related complex; chondrocyte growth; bone regeneration;	
KM	periodontal regeneration; tissue transport; bone graft; skin aging;	
KM	keratinocyte growth; hair loss; melanocyte growth; cell proliferation;	
KM	cell growth; organ transplant; cell differentiation; body height;	
KM	weight; hair colour; eye colour; skin; percentage of adipose tissue;	
KM	pigmentation; cosmetic surgery; metabolism; biophyhm; cardiac rhythm;	
KM	depression; tendency for violence; pain; reproductive capability;	
KM	hormone level; endocrine level; appetite; libido; memory; stress;	
KM	storage capability; fat content; lipid content; protein content;	
KM	carbohydrate content; vitamin content; cofactor content;	
XX	nutritional component.	
OS	Homo sapiens.	
XX		
US	US2002147140-A1.	
XX		
PD	10-OCT-2002.	
XX		
PF	17-JAN-2001; 2001US-0764877.	
XX		
PR	31-JAN-2000; 2000US-179065P.	
PR	04-FEB-2000; 2000US-180628P.	
PR	08-JUN-2000; 2000US-214886P.	
PR	07-JUL-2000; 2000US-216647P.	
PR	07-JUL-2000; 2000US-216880P.	
PR	11-JUL-2000; 2000US-217487P.	
PR	11-JUL-2000; 2000US-217496P.	
PR	14-JUL-2000; 2000US-218290P.	
PR	26-JUL-2000; 2000US-220963P.	
PR	26-JUL-2000; 2000US-220964P.	
PR	14-AUG-2000; 2000US-224518P.	
PR	14-AUG-2000; 2000US-224519P.	
PR	14-AUG-2000; 2000US-225267P.	

PR	14-AUG-2000;	2000US-225268P.
PR	14-AUG-2000;	2000US-225272P.
PR	14-AUG-2000;	2000US-225247P.
PR	14-AUG-2000;	2000US-225757P.
PR	14-AUG-2000;	2000US-225758P.
PR	22-AUG-2000;	2000US-226868P.
PR	30-AUG-2000;	2000US-228924P.
PR	01-SEP-2000;	2000US-229287P.
PR	01-SEP-2000;	2000US-229343P.
PR	01-SEP-2000;	2000US-229344P.
PR	01-SEP-2000;	2000US-229345P.
PR	05-SEP-2000;	2000US-229509P.
PR	05-SEP-2000;	2000US-229513P.
PR	08-SEP-2000;	2000US-231413P.
PR	21-SEP-2000;	2000US-234223P.
PR	21-SEP-2000;	2000US-234274P.
PR	25-SEP-2000;	2000US-234997P.
PR	27-SEP-2000;	2000US-235834P.
PR	29-SEP-2000;	2000US-236327P.
PR	29-SEP-2000;	2000US-236367P.
PR	29-SEP-2000;	2000US-236368P.
PR	29-SEP-2000;	2000US-236369P.
PR	29-SEP-2000;	2000US-236369P.
PR	29-SEP-2000;	2000US-236370P.
PR	02-OCT-2000;	2000US-236802P.
PR	02-OCT-2000;	2000US-237033P.
PR	02-OCT-2000;	2000US-237038P.
PR	02-OCT-2000;	2000US-237049P.
PR	02-OCT-2000;	2000US-237049P.
PR	13-OCT-2000;	2000US-239335P.
PR	20-OCT-2000;	2000US-240960P.
PR	20-OCT-2000;	2000US-241785P.
PR	20-OCT-2000;	2000US-241801P.
PR	01-NOV-2000;	2000US-244617P.
PR	17-NOV-2000;	2000US-249299P.
PR	08-DEC-2000;	2000US-251856P.
PR	08-DEC-2000;	2000US-251868P.
PR	08-DEC-2000;	2000US-251869P.
XX		
PA	(ROSE/) ROSEN C A.	
PA	(RUBE/) RUBEN S M.	
PA	(BARA/) BARASH S C.	
XX		
PI	Rosen CA, Ruben SM, Barash SC;	
XX		
DR	WPI; 2003-128199/12.	
DR	P-PSDB; ABU12769.	
XX		
PT	Isolated nucleic acid molecules encoding musculoskeletal system	
XX	associated polypeptides, useful for detecting disorders, e.g. cancer -	
XX		
PS	Claim 1; SEQ ID NO 399; 321p; English.	
XX		
CC	The invention describes an isolated nucleic acid molecule comprising a	
CC	sequence encoding musculoskeletal system associated polypeptides useful	
CC	for detecting disorders, e.g., cancer or cancer metastases, in animals	
CC	or humans. The nucleic acid; stimulates re-vascularisation of ischaemic	
CC	tissues associated with conditions such as thrombosis, arteriosclerosis,	
CC	and other cardiovascular conditions; treats wounds due to injuries,	
CC	burns, post-operative tissue repair, and ulcers; stimulates angiogenesis	
CC	and limb regeneration; stimulates neuronal growth; can treat and prevent	
CC	neuronal damage occurring in certain disorders or neurodegenerative	
CC	conditions, such as, Alzheimer's disease, Parkinson's disease, and	
CC	AlDS-related complex; stimulates chondrocyte growth, thus they can be	
CC	used to enhance bone and periodontal regeneration and aid in tissue	
CC	transports or bone grafts; prevents skin aging due to sunburn by	
CC	stimulating keratinocyte growth; prevents hair loss; since FGF family	
CC	members activate hair-forming cells and promotes melanocyte growth;	
CC	stimulates growth and differentiation of hematopoietic cells and bone	
CC	marrow cells when used in combination with other cytokines; maintains	
CC	organs before transplantation or for supporting cell culture of primary	
CC	tissues; induces tissue of mesodermal origin to differentiate in early	
CC	embryos; increases or decreases the differentiation or proliferation of	
CC	embryonic stem cells, besides, haematopoietic lineage; modulates	

CC mammalian characteristics, such as, body height, weight, hair colour, eye
CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
CC mammal's mental state or physical state by influencing biorhythms,
CC circadian rhythms, depression, tendency for violence, tolerance for pain,
CC reproductive capabilities, hormonal or endocrine levels, appetite,
CC libido, memory, or stress; increases or decreases storage capabilities,
CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
CC or other nutritional components. This sequence encodes a novel human
CC musculoskeletal system antigen.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.

XX
SQ Sequence 222 BP; 44 A; 63 C; 75 G; 38 T; 2 other;

Query Match 3.6%; Score 58; DB 25; Length 222;

Best local Similarity 100.0%; Pred. No. 1.1e-18; Mismatches 0; Gaps 0;

Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1521 CCACTGAGGAGTGTTCCTTCCCTTGTGCTTAAAGGCAAGAGCGAGCGGA 1578
|||
Db 32 CCACTGAGGAGTGTTCCTTCCCTTGTGCTTAAAGGCAAGAGCGAGCGGA 89

Search completed: November 25, 2003, 02:31:18
Job time : 500.846 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2003, 00:22:12 ; Search time 6309.68 Seconds
(without alignments)
10529.410 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 1624
Sequence: 1 gtcgggctgcgcgcgcgcaccc.....ccaggctctgtgcgggcccag 1624

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 15

Total number of hits satisfying chosen parameters: 99189

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank1:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1624	100.0	1624	6	AX449217
2	1304	80.3	2160	6	BD157613
3	1304	80.3	2160	6	AK022317
4	920	56.7	1022	9	AF237290
5	880	54.2	147492	9	AL358781
6	834	51.4	834	6	AX449219
7	783	48.2	834	9	AF254133
8	654	40.3	2072	9	BC015547
9	653	40.2	2228	9	AK057848
10	640	39.4	1395	9	AF125106
11	623	38.4	753	6	BD146824
12	338	20.8	734	6	AX540411
13	36	2.2	660	11	BY077668
14	36	2.2	1810	10	MUSURKI
15	36	2.2	1859	10	BC025146
16	36	2.2	233210	2	AC097693
17	36	2.2	241882	10	AL0808027
18	36	2.2	244105	2	AC098897
19	36	2.2	254577	2	AC068494
20	33	2.0	254993	2	AC078885
21	23	1.4	336	9	BT006860
22	22	1.4	336	12	BT007548
23	23	1.4	383	11	G41891
24	23	1.4	744	9	AB062451
25	23	1.4	1121	9	AF236637
26	23	1.4	1209	9	BC002906
27	23	1.4	1322	6	AX135546
28	23	1.4	99866	2	AL358115
29	23	1.4	164246	2	AC016371
30	23	1.4	178657	2	AC140447
31	22	1.4	841	11	BY017303
32	22	1.4	952	10	MUSURKIA
33	22	1.4	13852	9	AB026490
34	22	1.4	143291	9	HS16369
35	22	1.4	165248	2	AC102308
36	22	1.4	192336	2	AC118474
37	22	1.4	204529	2	AC021445
38	22	1.4	211075	10	AC076974
39	22	1.4	221847	2	AC127929
40	22	1.4	231524	2	AC115181
41	22	1.4	231337	2	AC127758
42	22	1.4	265928	2	AC093937
43	22	1.4	271603	2	AC106318
44	22	1.4	284499	2	AC134366
45	21	1.3	448	10	AF238445

ALIGNMENTS

RESULT 1
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Glucksmann, M. A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

FEATURES Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGEDCESPAPEADRPHPRLIGVSGTASGKSTCEKIM
ELIGONEVEORORRVILSODRFYKVLTAOKAKALGOVNPDPFNDLHRLTK
NIVEKTEVEPYDFVTHSRILPETHVYPADVLFEILVFGSOILRDMHILFPLDT
DSVRLSRVLRVRRGRDLEQILTYTTTKVPAFEFCCLPTKXIVADVITPREVDNV
AIIIVOHIDILNGDICMKHRRGSNRSYKRTFSEPDHPMLTSGKRLHSSRP
H"
BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN
Query Match 100.0%; Score 1624; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGAGTCCGCTCCGACCTCGAGCGCTGGAGCGGCGCGCGGGAAGGGGCGGCG
DB 1 GTGGAGTCCGCTCCGACCTCGAGCGCTGGAGCGGCGCGCGGGAAGGGGCGGCG
QY 61 GCGGGGACCCGATCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
DB 61 GCGGGGACCCGATCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
QY 121 GAGAGCCCCCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
DB 121 GAGAGCCCCCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
QY 121 GAGAGCCCCCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
DB 121 GAGAGCCCCCGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
QY 181 GCGGCACTGCGACGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
DB 181 GCGGCACTGCGACGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
QY 241 AACGAGGTGAAACGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
DB 241 AACGAGGTGAAACGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
QY 301 GTCTGACCGGAGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
DB 301 GTCTGACCGGAGCGGAGCGGAGCCGAGATGGCTTCGCGGAGGCGAGACTGC
QY 361 GATGCTTTGATATGATTTGATGACAGACTCTGAAGAACATCGTGAGGGCAAAACG
DB 361 GATGCTTTGATATGATTTGATGACAGACTCTGAAGAACATCGTGAGGGCAAAACG
QY 421 GTGAGAGTGCCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCTGTC
DB 421 GTGAGAGTGCCTATGATTTTGTGACACACTCAAGGTTACAGAGACACGCTGTC
QY 481 TACCGTCCGAGCGGAGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
DB 481 TACCGTCCGAGCGGAGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
QY 541 GACATGTTCCACTGCGCTCTTCTGTGAGACCGACTCCGACGTCAGGCTGTCTGAGA
DB 541 GACATGTTCCACTGCGCTCTTCTGTGAGACCGACTCCGACGTCAGGCTGTCTGAGA
QY 601 GTTCTCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
DB 601 GTTCTCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGATGGCTTCGCGGAGGCGAGACTGC
QY 661 TTGCTGAAGCGGCTTCGAGAGATTTCTGCTGCGGCAAAAGAGATGTCGATGATC
DB 661 TTGCTGAAGCGGCTTCGAGAGATTTCTGCTGCGGCAAAAGAGATGTCGATGATC
QY 721 ATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGAACATCCAGAC
DB 721 ATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGAACATCCAGAC

DB 721 ATCCCAAGAGAGTGAACAATATGTTGCCATCAACTGATCGTGAACATCCAGAC
QY 781 ATTCTGAATGTGATCACTTCGAAATGCGACCGAGAGGGTCCAAATGGCGAGCTCAAG
DB 781 ATTCTGAATGTGATCACTTCGAAATGCGACCGAGAGGGTCCAAATGGCGAGCTCAAG
QY 841 CGGACCTTTTCTGAGCGGAGGAGCCACCTTGGAGTGTCTGAGGCAAAAGGTCACAT
DB 841 CGGACCTTTTCTGAGCGGAGGAGCCACCTTGGAGTGTCTGAGGCAAAAGGTCACAT
QY 901 TTGAGTCCAGACGAGACCCCACTGAGGGGCTGCGGAGCTTCAGGGCTTCGCGGCG
DB 901 TTGAGTCCAGACGAGACCCCACTGAGGGGCTGCGGAGCTTCAGGGCTTCGCGGCG
QY 961 CGGACATGTGTTTCAAGGACTGAGCTGGGAGAGCCCAACCACTGCTTCTCTC
DB 961 CGGACATGTGTTTCAAGGACTGAGCTGGGAGAGCCCAACCACTGCTTCTCTC
QY 1021 GCGGCAACCCAGGAGGTGTTAGCAGCGAGGCTTCCTCACTCAGAGTGGAACTCAGA
DB 1021 GCGGCAACCCAGGAGGTGTTAGCAGCGAGGCTTCCTCACTCAGAGTGGAACTCAGA
QY 1081 TGTGTCACTCAAGACTCAACTTCTGAGGACACTGACAGGCTTCGAGGTTTTCAGCCAC
DB 1081 TGTGTCACTCAAGACTCAACTTCTGAGGACACTGACAGGCTTCGAGGTTTTCAGCCAC
QY 1141 TTAGGCTGTGTTGAGGTTTAAAGTCCCTCTAGGTCACTGAGAAATGCCAAGATGTGA
DB 1141 TTAGGCTGTGTTGAGGTTTAAAGTCCCTCTAGGTCACTGAGAAATGCCAAGATGTGA
QY 1201 GGAAGCTGAGAGGCTTCGTGAGAGATGTGAGAGCAATTAATGGGGAATGAGAGAGC
DB 1201 GGAAGCTGAGAGGCTTCGTGAGAGATGTGAGAGCAATTAATGGGGAATGAGAGAGC
QY 1261 AGCCTAGACACTGCTGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGAGTT
DB 1261 AGCCTAGACACTGCTGCTGATGTTTGTGACAGTGAACCAAGTGGAGAGAGTT
QY 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTACTCAAAAGTTTGTGAACA
DB 1321 TTTTCCAGTGTGATCTGTTCTTACACACTCAACATTACTCAAAAGTTTGTGAACA
QY 1381 AGTACTTCTTTTACATGTTACATGTCATGTTCTGTTTCTGTTTCTGTTTCAATA
DB 1381 AGTACTTCTTTTACATGTTACATGTCATGTTCTGTTTCTGTTTCTGTTTCAATA
QY 1441 CAAGGCTGTGCTGCTCAAAACCTAATTTATGACCCAGTGTGAGCTGCAAGCTG
DB 1441 CAAGGCTGTGCTGCTCAAAACCTAATTTATGACCCAGTGTGAGCTGCAAGCTG
QY 1501 GCTTACAGGATATGAGGAGCCACTGAGGAGATGTTTCCCTTGTGCTGCTTAAAG
DB 1501 GCTTACAGGATATGAGGAGCCACTGAGGAGATGTTTCCCTTGTGCTGCTTAAAG
QY 1561 GCAGAGAACGAGGCGGAGTCCCTGGAAGCACCCAGATCAACCCAGGCTGTGCGGGG
DB 1561 GCAGAGAACGAGGCGGAGTCCCTGGAAGCACCCAGATCAACCCAGGCTGTGCGGGG
QY 1621 CCAG 1624
DB 1621 CCAG 1624
RESULT 2
BD157613
LOCUS BD157613 2160 bp DNA linear PART 17-JAN-2003
DEFINITION primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613.1 GI:27863371
VERSION 1
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2160)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 12456 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/12456
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOHSHIO OTA,TAKAO ISOGAI,TERUSO NISHIKAWA,KOJI HAYASHI,KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
 PI KEIICHI NAGAI,TERUSUI OTSUKI
 PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
 10,
 PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key
 Location/Qualifiers
 FT CDS
 Location/Qualifiers
 1..2160
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 BASE COUNT 457 a 591 c 671 g 441 t
 Query Match 80.3%; Score 1304; DB 6; Length 2160;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 46 GGGGAAAGGGCGGCGCGGCGGAGCCCGATGCGCGGAGCGGAGCGGAGATGCGCTTCGGCG 105
 DB 47 GGGGAAAGGGCGGCGCGGCGGAGCCCGATGCGCGGAGCGGAGCGGAGATGCGCTTCGGCG 106
 QY 106 GGAAGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 165
 DB 107 GGAAGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 166
 QY 166 CTGATAGGGGCTGAGCGCGGCGGAGCTGCGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 225
 DB 167 CTGATAGGGGCTGAGCGCGGCGGAGCTGCGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 226
 QY 226 GAGTTCCTGGGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 285
 DB 227 GAGTTCCTGGGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGATGCGCTTCGGCG 286
 QY 286 GACAGGTTCTACAGGTTCTCTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
 DB 287 GACAGGTTCTACAGGTTCTCTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346
 QY 346 AATTTTGACATCCAGATGCTTTGATTAATGATTTGATGACAGGAGCTCTGAAGAGAGATC 405
 DB 347 AATTTTGACATCCAGATGCTTTGATTAATGATTTGATGACAGGAGCTCTGAAGAGAGATC 406
 QY 406 GTGGAGGAG 465
 DB 407 GTGGAGGAG 466
 QY 466 GAGACCAAGGAGGAG 525
 DB 467 GAGACCAAGGAGGAG 526
 QY 526 AGGACAG 585
 DB 527 AGGACAG 586
 QY 586 AGGCTCTCTGAG 645

DB 587 AGGCTCTCTGAG 646
 QY 646 AGGCTCTCTGAG 705
 DB 647 AGGCTCTCTGAG 706
 QY 706 TATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765
 DB 707 TATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 766
 QY 766 GAGACATCCAG 825
 DB 767 GAGACATCCAG 826
 QY 826 GGGCGGAG 885
 DB 827 GGGCGGAG 886
 QY 886 GGCAG 945
 DB 887 GGCAG 946
 QY 946 GGCAG 1005
 DB 947 GGCAG 1006
 QY 1006 CCACTGCT 1065
 DB 1007 CCACTGCT 1066
 QY 1066 GAGTGAAG 1125
 DB 1067 GAGTGAAG 1126
 QY 1126 GAGTGAAG 1185
 DB 1127 GAGTGAAG 1186
 QY 1186 GGCAG 1245
 DB 1187 GGCAG 1246
 QY 1246 GGAATTTGAG 1305
 DB 1247 GGAATTTGAG 1306
 QY 1306 CAGTGGAG 1349
 DB 1307 CAGTGGAG 1350
 RESULT 3
 AK022317 2160 bp mRNA linear PRI 01-AUG-2002
 LOCUS Homo sapiens cDNA FLJ12255 fis. clone MAMMA1001476, highly similar
 DEFINITION TO URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK022317
 VERSION AK022317.1 GI:10433687
 KEYWORDS oligo cloning; full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Shiraishi,A., Sudo,H.,
 Magatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
 Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakami,K.,
 Ono,Y., Takiguchi,S., Metanabe,S., Kimura,K., Murakami,K.,
 Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
 Nakamura,Y., Nagahara,K., Masuno,Y., Niinomiya,K. and Iwayanagi,T.
 NEDO human cDNA sequencing project
 TITLE Unpublished

REFERENCE 2 (bases 1 to 2160)
AUTHORS Isegai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isegai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5' - & 3' - end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1. .2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA1001476"
/tissue_type="Mammary gland"
/clone_id="MAMMA1"
/note="Cloning vector: pME18SFL3"
95. .928
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASGEBDESPAREADRPQRPFLIGVSGTASGKSTCEKIM
ELIGONEVEORRQKVVILISQDRFYKVLTAQKAKALKQYNFPHPDADFMDLHRLTK
NIVEGKTEVEPTDVFTHSRLEPTTVVYPADVVFEGILVFEQIEIDFNHLLFDYT
DSVRLSRVLRDVRGRDLLOILTOYTTVKPAFEFFCPTKRYADVILPRVDNVV
AIVLIVHIDLIANGDICKMHRGSSNRSYKRTFSEGDHPGLTSGKSHLESSSRP
H"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 80.3%; Score 1304; DB 9; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 GGGGAAAGGGGCGGCGGGGACCCGATGCGCGGAGCGAGCGGACGAGATGCTTCGGCG 105
DB 47 GGGGAAAGGGGCGGCGGGGACCCGATGCGCGGAGCGGAGCGGACGAGATGCTTCGGCG 106
QY 106 GAGAGCGGAAGACTCGGAGAGCCCGCGCGCGGAGCGGACCGTCCGACACAGCGGCCCTTC 165
DB 107 GAGAGCGGAAGACTCGGAGAGCCCGCGCGCGGAGCGGACCGTCCGACACAGCGGCCCTTC 166
QY 166 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGTGTGAAGATCATG 225
DB 167 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGTGTGAAGATCATG 226
QY 226 GAGTTGCTGGGACGAAAGAGGTGGAACAGCGGAGCGGAGGAGTGTATCTTGAGCCAG 285
DB 227 GAGTTGCTGGGACGAAAGAGGTGGAACAGCGGAGCGGAGGAGTGTATCTTGAGCCAG 286
QY 286 GACAGGTTCTCAAGAGTCTGACGCGAGAGTGAAGAGCCGAGGCTTGAAAGACAGTAC 345
DB 287 GACAGGTTCTCAAGAGTCTGACGCGAGAGTGAAGAGCCGAGGCTTGAAAGACAGTAC 346
QY 346 AATTGACCATTCAGATGCTTGTATGATGATTGTGATGACAGGACTCTGAAGAAATC 405
DB 347 AATTGACCATTCAGATGCTTGTATGATGATTGTGATGACAGGACTCTGAAGAAATC 406
QY 406 GTGAGAGGCAAAACGATGAGAGTCCGACCTATGATTGTCACACTCAAGTTTACA 465
DB 407 GTGAGAGGCAAAACGATGAGAGTCCGACCTATGATTGTCACACTCAAGTTTACA 466
QY 466 GAGACCAACGGTGTCTACCTCGCGAGAGTGTCTGTTTGAAGGCACTCTGTCTTAC 525
DB 467 GAGACCAACGGTGTCTACCTCGCGAGAGTGTCTGTTTGAAGGCACTCTGTCTTAC 526
QY 526 AGCCAGAGAGATCCGGGACATGTTCCACTGTGCGCTCTTCGTGACACCGACTCCGACGTC 585

DB 527 AGCCAGAGAGATCCGGGACATGTTCCACTGTGCGCTCTTCGTGACACCGACTCCGACGTC 586
QY 586 AGGCTGTCTCGAAGAGTCTCCGGGACGTGCCCGGAGGAGGACTGTGAGCATTTCTG 645
DB 587 AGGCTGTCTCGAAGAGTCTCCGGGACGTGCCCGGAGGAGGACTGTGAGCATTTCTG 646
QY 646 AGCAATTAACACACTTGTGAAAGCGGCGCTTGAGAGATTTGCTGCGCAAAAGAG 705
DB 647 AGCAATTAACACACTTGTGAAAGCGGCGCTTGAGAGATTTGCTGCGCAAAAGAG 706
QY 706 TATGCCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATCACTGATCTG 765
DB 707 TATGCCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATCACTGATCTG 766
QY 766 CAGCACATCCAGACATTTGTAATGTGATCATCTGCAAAATGCAACCGAGAGGTTCAAT 825
DB 767 CAGCACATCCAGACATTTGTAATGTGATCATCTGCAAAATGCAACCGAGAGGTTCAAT 826
QY 826 GGGCGGAGCTTCAAGGGAAGCTTTTCTGAGCCAGGAGGACCACTGGAGTGTGACTCT 885
DB 827 GGGCGGAGCTTCAAGGGAAGCTTTTCTGAGCCAGGAGGACCACTGGAGTGTGACTCT 886
QY 886 GGCAGACGCTCATATTGAGATCCAGACAGACACCCCATGAGGGGCTGCCAGCTCAG 945
DB 887 GGCAGACGCTCATATTGAGATCCAGACAGACACCCCATGAGGGGCTGCCAGCTCAG 946
QY 946 GGCAGGCTTCCGCGCGGAGTGTGTTGAGGACTGAGCTTGGGAGCGCCACCCAC 1005
DB 947 GGCAGGCTTCCGCGCGGAGTGTGTTGAGGACTGAGCTTGGGAGCGCCACCCAC 1006
QY 1006 CCACTGCTTCTCTCGGCGGACCCCGAGGAGTGTGAGAGGAGGCTTCTCACTCAG 1065
DB 1007 CCACTGCTTCTCTCGGCGGACCCCGAGGAGTGTGAGAGGAGGCTTCTCACTCAG 1066
QY 1066 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGCACTGACAGGCTTCT 1125
DB 1067 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGCACTGACAGGCTTCT 1126
QY 1126 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGCACTGACAGGCTTCT 1185
DB 1127 GAGTGGAAACTAGATGTGTCACTAGACTCAACTTGTGAGCACTGACAGGCTTCT 1186
QY 1186 GGCACAGATATGACAGAGAGCTTGGAGAGCTTCTGAGAGATGTGAGGCACTAATTG 1245
DB 1187 GGCACAGATATGACAGAGAGCTTGGAGAGCTTCTGAGAGATGTGAGGCACTAATTG 1246
QY 1246 GGAATTTGAGAGACAGCTTGAACACTGAGCTGAGTGTGTTGTTGACAGTGAACCA 1305
DB 1247 GGAATTTGAGAGACAGCTTGAACACTGAGCTGAGTGTGTTGTTGACAGTGAACCA 1306
QY 1306 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTGATCTTCAACAC 1349
DB 1307 CAGTGGAGAGAGTTTTCAGTCTGATCTGATCTGATCTTCAACAC 1350
RESULT 4
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.
DEFINITION ACCSSION
VERSION AF237290.1 GI:13506764
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Van Kompay, A.R., Norda, A., Linden, K., Johansson, M. and Karlsson, A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases
JOURNAL Mol. Pharmacol. 59 (5), 1181-1186 (2001)
MEDLINE 21203813

PUBMED 11306702
 REFERENCE 2 (bases 1 to 1022)
 AUTHORS Van Rompay, A.R., Lindén, K., Norda, A., Zhu, C., Zheng, X.,
 Johansson, M., and Karlsson, A.
 TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1022)
 AUTHORS Van Rompay, A.R., Lindén, K., Norda, A., Zhu, C., Zheng, X.,
 Johansson, M., and Karlsson, A.
 TITLE Direct Submission
 JOURNAL Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
 University Hospital, Stockholm 14186, Sweden
 FEATURES Location/Qualifiers
 source 1..1022
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..1022
 /gene="UCK1"
 5..838
 /gene="UCK1"
 /function="phosphorylates uridine and cytidine to uridine
 monophosphate and cytidine monophosphate"
 /codon_start=1
 /product="uridine-cytidine kinase 1"
 /protein_id="AAK28324.1"
 /db_xref="GI:13506765"
 /translation="MASAGEDESPAPREAPRPHRPFILVSGTASGKSTCEKIM
 ELGQNEVERQKRVVILSDREYKVLTAQKAKALKGQNFPHAPFNDLHRTLK
 NIVGEKTEVPTVTHSLRPETVYVADYVLFGLIVFVQSLRDMFLHLPYDT
 DSPVRLSRVLRVRRGRDLLEOILTOYTFVKKRPAEPFCLPTKKYADVLIIPRDVMV
 AINLIYQHIQDILINQDIQKMRGSGNGRSYKRTFSRPHQMLTSGKSHLESSRP
 H"

BASE COUNT 227 a 282 c 324 g 189 t
 ORIGIN

Query Match 56.7%; Score 920; DB 9; Length 1022;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1020; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 90 CGAGATGGCTTCGCGGGAGCGAAGACTGCGAGAGCCCGCGGAGCCGACCGTCC 149
 DB 1 CGAGATGGCTTCGCGGGAGCGAAGACTGCGAGAGCCCGCGGAGCCGACCGTCC 60
 QY 150 GCAACCGCGGCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAATGCCGT 209
 DB 61 GCAACCGCGGCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAATGCCGT 120
 QY 210 GTGTGAAGATCATGAGTGTGCTGGGACGAACGAGGTGGAACAGCGCGAGCGGAAGT 269
 DB 121 GTGTGAAGATCATGAGTGTGCTGGGACGAACGAGGTGGAACAGCGCGAGCGGAAGT 180
 QY 270 GGTTCATCTGAGCGAGGACAGTTCCTCAAGTCTCTGACGCGAGAGCAAGGCCAAGGC 329
 DB 181 GGTTCATCTGAGCGAGGACAGTTCCTCAAGTCTCTGACGCGAGAGCAAGGCCAAGGC 240
 QY 330 CTTGAAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGACAG 389
 DB 241 CTTGAAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGACAG 300
 QY 390 GACTGTGAAGACATGCTGAGGAGGCAAAACGAGGAGGTGCGACCTATGATTTGAC 449
 DB 301 GACTGTGAAGACATGCTGAGGAGGCAAAACGAGGAGGTGCGACCTATGATTTGAC 360
 QY 450 ACACTGAAGGTTACAGAGACCAAGTGTCTACCCCTGCGAGCTGATTTGAGG 509
 DB 361 ACACTGAAGGTTACAGAGACCAAGTGTCTACCCCTGCGAGCTGATTTGAGG 420
 QY 510 CATCTGGTGTCTTACAGCGAGAGATCCGGGACATTTTCACTGGCGCTTCTTGGGA 569
 DB 421 CATCTGGTGTCTTACAGCGAGAGATCCGGGACATTTTCACTGGCGCTTCTTGGGA 480
 QY 570 CACCGACTCGAGCTGCTGTGGAAGAGTTTCCCGGAGCTGCGCGGAGGAGGA 629

DB 481 CACCGACTCGAGCTGCTGTGGAAGAGTTTCCGGGAGCTGCGCGGAGGAGGA 540
 QY 630 CTTGAGACGATTTGACGACAGTACCACTTCTGTAAGCGGCGCTTGGAGAGTTCTG 689
 DB 541 CTTGAGACGATTTGACGACAGTACCACTTCTGTAAGCGGCGCTTGGAGAGTTCTG 600
 QY 690 CTTGCGGAGAAAGATAGCCGATGTGATCCACAGAGAGTGAACAATATGTTCC 749
 DB 601 CTTGCGGAGAAAGATAGCCGATGTGATCCACAGAGAGTGAACAATATGTTCC 660
 QY 750 CATCAACTGATCTGTCAGACACATCCAGAGACATTCTGATGTGACATCTGCAAAATGGA 809
 DB 661 CATCAACTGATCTGTCAGACACATCCAGAGACATTCTGATGTGACATCTGCAAAATGGA 720
 QY 810 CCGAGAGAGGTCATATGCGGAGGACTCAAGCGGACCTTTCTGAGCGAGGACCAACC 869
 DB 721 CCGAGAGAGGTCATATGCGGAGGACTCAAGCGGACCTTTCTGAGCGAGGACCAACC 780
 QY 870 TGGGATCTGACCTCTGCGAAACGGTCAATTGAGTGCACAGAGACCCCACTGAG 929
 DB 781 TGGGATCTGACCTCTGCGAAACGGTCAATTGAGTGCACAGAGACCCCACTGAG 840
 QY 930 GGCTGCGAGCTTCAGGGGAGGTTCCCGCGGACATGTGTTCAGAGGACTGAGCTGG 989
 DB 841 GGCTGCGAGCTTCAGGGGAGGTTCCCGCGGACATGTGTTCAGAGGACTGAGCTGG 900
 QY 990 GGAACGCCACCAACCACTCTCTCTCTGCGGACCCAGGAGGAGTGTAGACGA 1049
 DB 901 GGAACGCCACCAACCACTCTCTCTCTGCGGACCCAGGAGGAGTGTAGACGA 960
 QY 1050 GGCTTCTCTCACTGAGATGGAATCTGAGTGTGATCTGACATCTGCTGGGAC 1109
 DB 961 GGCTTCTCTCACTGAGATGGAATCTGAGTGTGATCTGACATCTGCTGGGAC 1020
 QY 1110 AC 1111
 DB 1021 AC 1022

RESULT 5
 AL358781/c
 LOCUS AL358781 147492 bp DNA linear PRI 06-OCT-2001
 DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete
 sequence.
 ACCESSION AL358781
 VERSION AL358781.19 GI:13751418
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 147492)
 REQUESTER: clonerequest@sanger.ac.uk
 On Apr 21, 2001 this sequence version replaced gi:13396549.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given

QY 454 TCAGGTTACAGAGACCAACGATGATCTACCTGCGGAGTGGTCTGTTGAGGGGATC 513
Db 361 TCAGGTTACAGAGACCAACGATGATCTACCTGCGGAGTGGTCTGTTGAGGGGATC 420
QY 514 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 573
Db 421 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 480
QY 574 GATTCGACGTCAGGCTGTCTCCAGAGATCTCCGGGACGTCGCGGAGGAGGACCTG 633
Db 481 GATTCGACGTCAGGCTGTCTCCAGAGATCTCCGGGACGTCGCGGAGGAGGACCTG 540
QY 634 GACGAGATTCGACGAGTACACACCTTGTGTAAGCGGCTTCCAGAGATCTGCTG 693
Db 541 GACGAGATTCGACGAGTACACACCTTGTGTAAGCGGCTTCCAGAGATCTGCTG 600
QY 694 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 753
Db 601 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 660
QY 754 AACCTGATCTGACGACATCCAGGACATCTGTAATGTCATCTGCAATATGACACCA 813
Db 661 AACCTGATCTGACGACATCCAGGACATCTGTAATGTCATCTGCAATATGACACCA 720
QY 814 GGAGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 873
Db 721 GGAGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 780
QY 874 ATCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGACCCCACTGA 927
Db 781 ATCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGACCCCACTGA 834

RESULT 7
LOCUS AF254133 834 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
Human uridine kinase from prostate cancer cell line (LNCap)
Unpublished
2 (bases 1 to 834)
Ho, Y.S. and Johnson, R.K.
Direct Submission
Submitted (10-APR-2000) Oncology Research, Smlthkline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA
FEATURES
source
1..834
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAK49122.1"
/db_xref="GI:13924750"
translation="MASAGGECDESPAPEDRPHORPFLIGVSGTASGKSTVCERIM
ELIGNEVEORORRVVILSGDRYKVLTAOKAKLKGQVNFDPHAFDNLHRTLK
NIVEGKTEVEPTDFTVTHSRLEPTTVVYPADVVLFGSLIVYSOELRDMFHLRLFVDT
DSPVRLSRVLRDVRGRDLEQLITQYTTTVKRAPEFCLPTKYADVITPRVDMMV
AIVLIVQIIDIINGDI CKMRGSGNSRSTKRTFSEPDHPMLTSGKSHLESSSRP
H"

BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN

Query Match 48.2%; Score 783; DB 9; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 94 ATGGCTTCGCGGAGGCGAGAGACTGCGAGACCCCGCGCGGAGGCGGACCGTCCGAC 153
Db 1 ATGGCTTCGCGGAGGCGAGAGACTGCGAGACCCCGCGCGGAGGCGGACCGTCCGAC 60
QY 154 CAGCGGCCCTTCTCTGATATAGGGGTGAGCGCGGACCTGCGAGGGGAAAGTGCACCTGTGT 213
Db 61 CAGCGGCCCTTCTCTGATATAGGGGTGAGCGCGGACCTGCGAGGGGAAAGTGCACCTGTGT 120
QY 214 GAGAAATCATGAGAGTGTCTGAGGACAGAACTGAGTGGAAACAGCGGACGGAAGGTGTC 273
Db 121 GAGAAATCATGAGAGTGTCTGAGGACAGAACTGAGTGGAAACAGCGGACGGAAGGTGTC 180
QY 274 ATCTGAGCCAGAGACAGGTTCTAACAAGTCTGACGCGGACAGAGGCAAGGCTTTG 333
Db 181 ATCTGAGCCAGAGACAGGTTCTAACAAGTCTGACGCGGACAGAGGCAAGGCTTTG 240
QY 334 AAAGGACATGACATTTTGAACCATCCAGATGCTTGTGATATATGATTTGACACAGACT 393
Db 241 AAAGGACATGACATTTTGAACCATCCAGATGCTTGTGATATATGATTTGACACAGACT 300
QY 394 CTGAAGAATCATGTGGAGGCAAAACGGTGAAGTGCAGGACCTATGATTTTGTGACACAC 453
Db 301 CTGAAGAATCATGTGGAGGCAAAACGGTGAAGTGCAGGACCTATGATTTTGTGACACAC 360
QY 454 TCAGGTTACAGAGACCAACGATGATCTACCTGCGGAGTGGTCTGTTGAGGGGATC 513
Db 361 TCAGGTTACAGAGACCAACGATGATCTACCTGCGGAGTGGTCTGTTGAGGGGATC 420
QY 514 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 573
Db 421 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTCGACAC 480
QY 574 GACTCCGAGTCAAGGCTGTCTGGAAGATTCCTCGGAGAGTGCAGGAGGAGGACCTG 633
Db 481 GACTCCGAGTCAAGGCTGTCTGGAAGATTCCTCGGAGAGTGCAGGAGGAGGACCTG 540
QY 634 GAGCAGATTCGACGAGTACACACCTTCTGTAAGCCCGGCTTCCAGAGATTTCTGCTG 693
Db 541 GAGCAGATTCGACGAGTACACACCTTCTGTAAGCCCGGCTTCCAGAGATTTCTGCTG 600
QY 694 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 753
Db 601 CCGACAAAGAGATGATCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 660
QY 754 AACCTGATCTGACGACATCCAGGACATCTGTAATGTCATCTGCAATATGACACCA 813
Db 661 AACCTGATCTGACGACATCCAGGACATCTGTAATGTCATCTGCAATATGACACCA 720
QY 814 GGAGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 873
Db 721 GGAGGGTCCAAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGGAGGACCACTCTGG 780
QY 874 ATCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGACCCCACTGA 927
Db 781 ATCTGACCTCTGCGCAACCGTCAATTTGAGTCCAGCAGACCCCACTGA 834

RESULT 8
LOCUS BC015547 2072 bp mRNA linear PRI 04-OCT-2001
DEFINITION Homo sapiens, Similar to uridine-cytidine kinase 1, clone MGC:9668
IMAGE:3845821, mRNA, complete cds.
ACCESSION BC015547
VERSION BC015547.1 GI:15930229
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2072)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxil.stanford.edu
 (Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.)
 Cloning distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAK Plate: 20 Row: f Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.
 Location/Qualifiers
 1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NIH MGC_65"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 87..692
 /codon_start=1
 /product="Similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGGDCSPPEADRPORPFLGVSGGTSGKSTCEKIM ELISQNEVQRKRVILISODRFYKVLAEKAKALGVNPFDPDAFDDMLHRTLK NIVGKTVETPTVTHSRLEPTTVYPADVVFESILVFEYSQELRDMFHLRLFVDT DSDVLRSRDREVCRCCHPTPSGQYGHOPBRAHPSHSEW"
 BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN
 Query Match 40.3%; Score 654; DB 9; Length 2072;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 696 GACAAAGAGTATGCGGATGATCATCCCGAGAGGTGACAAATGTTGCATCAA 755
 Db 594 GACAAAGAGTATGCGGATGATCATCCCGAGAGGTGACAAATGTTGCATCAA 653
 QY 756 CCTGATGTCGACGACATCCAGCAATTCGAATGTCACATCTGCAATGGACCGAG 815
 Db 654 CCTGATGTCGACGACATCCAGCAATTCGAATGTCACATCTGCAATGGACCGAG 713
 QY 816 AGGATCCAAATGGCGAGACTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGGAT 875
 Db 714 AGGATCCAAATGGCGAGACTCAAGCGGACCTTTTCTGAGCCAGGGGACCACTTGGGAT 773
 QY 876 GCTGACCTTGGGAAAGGTGACATTTGGAGTCCAGAGGACCCCACTGAGGGGGCTGC 935
 Db 774 GCTGACCTTGGGAAAGGTGACATTTGGAGTCCAGAGGACCCCACTGAGGGGGCTGC 833
 QY 936 CGAGCTTCAGGGGAGTCTCCCGCCGAGCATGTGTTCAGAGGAGCTGAGCTGGGAGCG 995
 Db 834 CGAGCTTCAGGGGAGTCTCCCGCCGAGCATGTGTTCAGAGGAGCTGAGCTGGGAGCG 893
 QY 996 CGAGCTTCAGGGGAGTCTCCCGCCGAGCATGTGTTCAGAGGAGCTGAGCTGGGAGCG 1055

Db 894 CCAACCCACACCACTGCTTCTCTCGGCGCAGCCCGAGGAGTGTAGCAGGAGCCTT 953
 QY 1056 CTTCACTCAGAGTGGAACTCAGATGTGACATCCCACTTCTGGGACACTGAC 1115
 Db 954 CTTCACTCAGAGTGGAACTCAGATGTGACATCCCACTTCTGGGACACTGAC 1013
 QY 1116 AGGCGTTCCTGAGGTTTTCAGCCACTTAGGCTGCTGCGGTTTAAAGATCCCTAGTGC 1175
 Db 1014 AGGCGTTCCTGAGGTTTTCAGCCACTTAGGCTGCTGCGGTTTAAAGATCCCTAGTGC 1073
 QY 1176 ACTGAGAAATGCCACAGATGTGAGAGAGCTTGGAGAGCTTGTGAGAAATGTGAGGC 1235
 Db 1074 ACTGAGAAATGCCACAGATGTGAGAGAGCTTGGAGAGCTTGTGAGAAATGTGAGGC 1133
 QY 1236 AACTTATTTGGGAAATTTGAGAGACAGCCCTAGACACTGAGCTGAGCTGATGTTTGTGAC 1295
 Db 1134 AACTTATTTGGGAAATTTGAGAGACAGCCCTAGACACTGAGCTGAGCTGATGTTTGTGAC 1193
 QY 1296 AGTGAACCCACAGTGGAGAGAGTCTTTCAGTCTGATCTGATCTTACACAC 1349
 Db 1194 AGTGAACCCACAGTGGAGAGAGTCTTTCAGTCTGATCTGATCTTACACAC 1247
 RESULT 9
 AK057848
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to URIDINE KINASE (EC 2.7.1.48).
 ACCESSION AK057848
 VERSION AK057848.1 GI:16553809
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiya, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Negai, K., Isogai, T., and Sugano, S.
 MEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2228)
 Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.
 Location/Qualifiers
 1..2228
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CBR05878"
 /tissue_type="brain"
 /clone_lib="CBR"
 /note="cloning vector: pME185FL3"
 BASE COUNT 478 a 598 c 696 g 456 t
 ORIGIN
 Query Match 40.2%; Score 653; DB 9; Length 2228;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 653; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY		697	ACAAAGAAGTATGCGCATGTGATCACTCCACGAGGAGTGGAACAATATGGTTGCCATCAAC	756
Db		768	ACAAAAGAAGTATGCGCATGTGATCACTCCACGAGGAGTGGAACAATATGGTTGCCATCAAC	827
OY		757	CTGATTCGTGCAGCACAATCATCAGAGCAATTCTGTAATGTGTGACATCTGCAATATGGCACCGAAGA	816
Db		828	CTGATTCGTGCAGCACAATCATCAGAGCAATTCTGTAATGTGTGACATCTGCAATATGGCACCGAAGA	887
OY		817	GAGTCCAAATGAGCGCGAGCTCAACAGCGGACCTTTTCTGAGCCAGGGGACACACTTGAGATG	876
Db		888	GAGTCCAAATGAGCGCGAGCTCAACAGCGGACCTTTTCTGAGCCAGGGGACACACTTGAGATG	947
OY		877	CTGACCTCTGGCAAACGGTCAATTGTGAGATCCAGACAGACCCTGAGAGGGCTGCC	936
Db		948	CTGACCTCTGGCAAACGGTCAATTGTGAGATCCAGACAGACCCTGAGAGGGCTGCC	1007
OY		927	GAGCCTCAGGGGAGGCTCTCCGCCCGCGGATGTGTTCAGGGAGCTGAGCCTGGGGAGGCC	996
Db		1008	GAGCCTCAGGGGAGGCTCTCTCCGCCCGCGGATGTGTTCAGGGAGCTGAGCCTGGGGAGGCC	1067
OY		997	CACCCACACCCAATGCTCTCTCTCGGCGCACCCCAAGGGAGTGTTAGCAGGAGGCTCTTC	1056
Db		1068	CACCCACACCCAATGCTCTCTCTCGGCGCACCCCAAGGGAGTGTTAGCAGGAGGCTCTTC	1127
OY		1057	CTCACTCAGAGTGGAAATCTAGATGTGTCACTAGACTCAACTTGTGGGACACTGACA	1118
Db		1128	CTCACTCAGAGTGGAAATCTAGATGTGTCACTAGACTCAACTTGTGGGACACTGACA	1187
OY		1117	GGCGTCTCTGAGGTTTTTCAAGCCACTTAGGCTGTGGGGTTTTTAAAGATCCCTTAGAGTCA	1176
Db		1188	GGCGTCTCTGAGGTTTTTCAAGCCACTTAGGCTGTGGGGTTTTTAAAGATCCCTTAGAGTCA	1247
OY		1177	CTGAGAAATGGCCACAGAAATGTGTGAGGAAGCCTGGAGGCTTCTGTGAGAAATGTGAGGCA	1238
Db		1248	CTGAGAAATGGCCACAGAAATGTGTGAGGAAGCCTGGAGGCTTCTGTGAGAAATGTGAGGCA	1307
OY		1237	CATTATTTGGGAAATTGTGAGGAGACAGCCTTAGACACTGGCTGGCTGATGTGTTTGTGACA	1296
Db		1308	CATTATTTGGGAAATTGTGAGGAGACAGCCTTAGACACTGGCTGGCTGATGTGTTTGTGACA	1367
OY		1297	GTCGAACCCACAGTGGGAGAGATTTTTTCCAGTCTGATCTGTGTTTCAACAC	1349
Db		1368	GTCGAACCCACAGTGGGAGAGATTTTTTCCAGTCTGATCTGTGTTTCAACAC	1420
RESULT 10				
LOCUS	AF125106		1395 bp	mRNA linear PRI 07-FEB-2002
DEFINITION	Homo sapiens uridine kinase mRNA, complete cds.			
ACCESSION	AF125106			
VERSION	AF125106.1	GI:18568108		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1395)			
AUTHORS	Xin,Y.R., Yu,L. and Zhao,S.Y.			
TITLE	Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1395)			
AUTHORS	Ding,J.B., Yu,L. and Zhao,S.Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433, People's Republic of China			
FEATURES	Location/Qualifiers			
source	1..1395			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			

Query Match	Similarity	Score	DB	Length
Beet Local Match	99.6%	640	9	1395
Matches	910	Conservative	0	Mismatches 3; Indels 1; Gaps 1
BASE COUNT	325 a	352 c	411 g	307 t
ORIGIN	<pre> 8..838 /codon_start=1 /product="uridine_kinase" /protein_id="PAL75943.1" /db_xref="GI:18568109" /translation="MASAGGEGAGAGANRPHORPELIVSGSGTASGKSTICEKIME LLEGNVEIORHKGAVILISODPEFYKVLRAEOKAKLKGQYNDHPDPAFNDMLMRLLTKR IYVEKTVIEPTVDPVTSRRLPEITVVYPADVLEPEGLIFVSOELRDMFHLRLPDTT SDVRLSRHVELADVRGRDLEOILTQYTTTFKPAEERFELPKTKADVTIPIGVDMNV INLIVQHIQDILNDGICMKMRHGSNGRTKTFSEPDHPMLTISGRKSHLESSRPI </pre>			
265	AAGGTGGTCACTCTGAGCCGAGACAGAGTTCTTACAAAGTCCCTGACGACAGACAGAAAGCC	324		
Db	176	AAGGTGGTCACTCTGAGCCGAGACAGAGTTCTTACAAAGTCCCTGACGACAGACAGAAAGCC	235	
Qy	325	AAGGCTTGAAGAAGACAGTACAAATTTTGAACATCCAGATGCGCTTGTATGATTTGATG	384	
Db	236	AAGGCTTGAAGAAGACAGTACAAATTTTGAACATCCAGATGCGCTTGTATGATTTGATG	295	
Qy	385	CACAGGACTCTGAAGAACATCGTGAAGGCGAAACCGGTGAGGTGCGCACTTATGATTTT	444	
Db	296	CACAGGACTCTGAAGAACATCGTGAAGGCGAAACCGGTGAGGTGCGCACTTATGATTTT	355	
Qy	445	GTGACACACTCAAGTTTACAGAACCCACGGTGGTCTTACCTGGGGAAGTGGTCTGTTT	504	
Db	356	GTGACACACTCAAGTTTACAGAACCCACGGTGGTCTTACCTGGGGAAGTGGTCTGTTT	415	
Qy	505	GAGGGCACTCTGGGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGGGCGCTCTTC	564	
Db	416	GAGGGCACTCTGGGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGGGCGCTCTTC	475	
Qy	565	GTGACACCCGACTCCGACGTCAGGCTGTCTTGAAGATTTCTCCGGGACGTGCGCCGAGG	624	
Db	476	GTGACACCCGACTCCGACGTCAGGCTGTCTTGAAGATTTCTCCGGGACGTGCGCCGAGG	535	
Qy	625	AGGGAAGTCTGAGACAGATTTCTGACGAGTACACCACTTCTGGAAGCCGGCTTGGAGAG	684	
Db	536	AGGGAAGTCTGAGACAGATTTCTGACGAGTACACCACTTCTGGAAGCCGGCTTGGAGAG	595	
Qy	685	TTTCGCTCCCGAAGAAAGATATGCGGATGTCATCCCAAGAGAGTGAACAAATATG	744	
Db	596	TTTCGCTCCCGAAGAAAGATATGCGGATGTCATCCCAAGAGAGTGAACAAATATG	655	
Qy	745	GTTGCCATCAACTGATTCGTGACACATCCAGAGCAATTCGAAATGTGACATCTGCAAA	804	
Db	656	GTTGCCATCAACTGATTCGTGACACATCCAGAGCAATTCGAAATGTGACATCTGCAAA	715	
Qy	805	TGGCACCGAGAGGGGTCCATGGGCGGAGCTTAAAGCGGACCTTTTCTGACCCAGGGAG	864	
Db	716	TGGCACCGAGAGGGGTCCATGGGCGGAGCTTAAAGCGGACCTTTTCTGACCCAGGGAG	775	
Qy	865	CACCTTGGAGTCTGACCTCTGCGCAAAAGGTGACATTTGAGAGTCCAGACGACACCCAC	924	
Db	776	CACCTTGGAGTCTGACCTCTGCGCAAAAGGTGACATTTGAGAGTCCAGACGACACCCAC	835	
Qy	925	TGAGGGGGTGCGAGAGCTCAGGGGAGGCTCTCCGCGCGGACATGTGTGTCAGGAGCTGAG	984	
Db	836	TGAGGGGGTGCGAGAGCTCAGGGGAGGCTCTCCGCGCGGACATGTGTGTCAGGAGCTGAG	895	
Qy	985	CCTGGGGAGCGCCACCCACACCACTGCTTCTTCTGCGCGACCCACAGGGAGTGTAGC	1044	
Db	896	CCTGGGGAGCGCCACCCACACCACTGCTTCTTCTGCGCGACCCACAGGGAGTGTAGC	955	
Qy	1045	AGCGAGGCTTCTCTCACTCAGAGAGTGAACATCAGATGTGTGACCTCAGACTCACTTGCT	1104	
Db	956	AGCGAGGCTTCTCTCACTCAGAGAGTGAACATCAGATGTGTGACCTCAGACTCACTTGCT	1014	

```

QY 1105 GGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTAGCGTCTTGCGTTAAAGAT 1164
D 1015 GGACACTGACAGGCGTTCTGAGGTTTTCAGCCACTTAGCGTCTTGCGTTAAAGAT 1074
QY 1165 CCCTCTAGTCACT 1178
D 1075 CCCTCTAGTCACT 1088

RESULT 11
BD146824 753 bp DNA linear PAT 17-JAN-2003
LOCUS BD146824
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD146824.1 GI:27852582
VERSION JP 2002191363-A/1667.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
Otsu, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1667 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/1667
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC
10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..753
Location/Qualifiers
FEATURES
source
1..753
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 158 a 196 c 261 g 135 t 3 others
ORIGIN
Query Match 38.4%; Score 623; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 623; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGGAGAGGGGGCGGGCGGGGACCCGATGCGCGGAGCGGAGCGCGGATGCTTCGCGC 105
D 47 GGGGAGAGGGGGCGGGCGGGGACCCGATGCGCGGAGCGGAGCGCGGATGCTTCGCGC 106
QY 106 GGAAGCCGAAGACTCCGAGAGCCCGCGCGCGGAGCCGACCGTCCGACACAGCGGCCCTTC 165
D 107 GGAAGCCGAAGACTCCGAGAGCCCGCGCGCGGAGCCGACCGTCCGACACAGCGGCCCTTC 166
QY 166 CTGATAGGGGGTGAAGCGGCGGCACTGCGAGCGGGAAGTGAACCGGTGGAAGATATATG 225
D 167 CTGATAGGGGGTGAAGCGGCGGCACTGCGAGCGGGAAGTGAACCGGTGGAAGATATATG 226
QY 226 GAGTGTCTGGGACGAAGCAAGGTGGAACAGCGGCGGGAAGGTGTCATCTTGAAGCAG 285
D 227 GAGTGTCTGGGACGAAGCAAGGTGGAACAGCGGCGGGAAGGTGTCATCTTGAAGCAG 286
QY 286 GACAGGTTCTCAAGGTCCTGACGGCGAGGCAAGAGGCCAAGGCTTGAAGGACAGTAC 345

```

```

D 287 GACAGGTTCTCAAGGTCCTGACGGAGAGCAGAAAGGCCAAGGCTTGAAGGACAGTAC 346
QY 346 AATTTGACATCGAGATGCTTGAATGATTTGATGACAGAGCTCGAAGACATC 405
D 347 AATTTGACATCGAGATGCTTGAATGATTTGATGACAGAGCTCGAAGACATC 406
QY 406 GTGAGAGGCAAAAAGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGATTACCA 465
D 407 GTGAGAGGCAAAAAGGTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGATTACCA 466
QY 466 GAGACCAAGGTGCTACCTCGGACAGTGTCTGTTTGAAGGCACTTGTGTTCTAC 525
D 467 GAGACCAAGGTGCTACCTCGGACAGTGTCTGTTTGAAGGCACTTGTGTTCTAC 526
QY 526 AGCAGAGATCCGGGACATGTTCCACTCGGCTCTGTTGAGACCCGACCTCGAGCTC 585
D 527 AGCAGAGATCCGGGACATGTTCCACTCGGCTCTGTTGAGACCCGACCTCGAGCTC 586
QY 586 AGGCTGTCTGGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTGAGCAGATTCTG 645
D 587 AGGCTGTCTGGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTGAGCAGATTCTG 646
QY 646 ACGCAGTACACCACTTCTGTGAA 668
D 647 ACGCAGTACACCACTTCTGTGAA 669

RESULT 12
AX540411 734 bp DNA linear PAT 23-NOV-2002
LOCUS AX540411
DEFINITION Sequence 23 from Patent WO02055738.
ACCESSION AX540411
VERSION AX540411.1 GI:25273437
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Panzer, S.R., Lincoln, S.E., Altus, C.M., Dufour, G.E., Hillman, J.L.,
Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Daffo, A.,
Marwaha, R., Chen, A.J., Chang, S.C., Gerstl, E.H., Peralela, C.H.,
David, M.H. and Lewis, S.A.
Molecules for disease detection and treatment
Patent: WO 02055738-A 23 18-JUL-2002;
Incyte Genomics, Inc. (US)
Location/Qualifiers
FEATURES
source
1..734
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/not_e="Incyte ID No: LI:215557.12:2001JAN12"
BASE COUNT 166 a 181 c 226 g 154 t 7 others
ORIGIN
Query Match 20.8%; Score 338; DB 6; Length 734;
Best Local Similarity 100.0%; Pred. No. 2; 6e-165;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 359 CAGATCCCTTTGATTAATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGGGCAAAA 418
D 12 CAGATCCCTTTGATTAATGATTTGATGACAGAGCTCTGAAGAAATCTGTGAGGGCAAAA 419
QY 419 CGGTGAGGTGCGGACCTTGAATTTGTGACACACTCAAGGTGTAACAGAGACCAAGGTG 478
D 72 CGGTGAGGTGCGGACCTTGAATTTGTGACACACTCAAGGTGTAACAGAGACCAAGGTG 479
QY 72 TCAACCTCGGAGCGGTGTTCTGTTGAGGAGATCTTGTGTTCTCAACCCAGAGATCC 538
D 132 TCAACCTCGGAGCGGTGTTCTGTTGAGGAGATCTTGTGTTCTCAACCCAGAGATCC 539
QY 539 GGGACATGTTCCACTGCGCTCTTCTGTGAGACCGAGCTTCGACGTGAGGCTGTCTGAA 598

```

Db 192 GGGACATGTTCCAGCTGCGCCTCTTCTGTCGACACCGACCTCCGACGTGAGGCTGTCTCGAA 251
Qy 599 GAGTTCCTCCGGGACGTGCGCGCGGAGGAGGACCTGGAGCATTTGACCGCAGTACCA 658
Db 252 GAGTTCCTCCGGGACGTGCGCGCGGAGGAGGACCTGGAGCATTTGACCGCAGTACCA 311
Qy 659 CTTTCGAGACCGGCTTCGAGAGTTCGCTGCGG 696
Db 312 CTTTCGAGACCGGCTTCGAGAGTTCGCTGCGG 349

RESULT 13
BV077668/c
LOCUS 660 bp DNA linear STS 31-MAY-2003
DEFINITION S21286227FDB.T0 CZECHII/E1 Mus musculus STS genomic, sequence
TAGGED SITE.
BV077668
BV077668.1 GI:31193463
STX.
Mus musculus (house mouse)
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
The mosaic structure of variation in the laboratory mouse genome
Nature 420 (6915), 574-578 (2002)
22354684
12466852
COMMENT

TITLE
JOURNAL
MEDLINE
PUBMED
12466852

CONTACT: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
Fax: 6172580903
Email: kersti@genome.wi.mit.edu
Primer A: None
Primer B: None
STS size: 660
STS size: 660
Protocol:
WGS-discovery: Paired-end low-coverage whole genome shotgun reads
were generated from 129S1/svmt, C3H/HeJ, and BALB/cByJ. The WGS
reads were placed uniquely on the MGS03 C57BL/6J assembly and SNP
detection was carried out by SSAHA-SNP. 225,000 reads were
annotated
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J
and the strain from which the particular read came. The validation
rate for these SNPs was estimated at approximately 98%.

FEATURES
source
1. .660
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CZECHII/E1"
/db_xref="taxon:10090"
/map="2.22-615 32611326-32611918"
/clone_lib="CZECHII/E1"
<1. .>660
174 a 196 c 174 g 116 t

STS
BASE COUNT 174 a 196 c 174 g 116 t
ORIGIN

Query Match 2.2%; Score 36; DB 11; Length 660;
Best Local Similarity 100.0%; Pred. No. 6.9e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 CCAGAGATCCGGACATGTTCCACCTGCGCCTCTT 563
Db 435 CCAGAGATCCGGACATGTTCCACCTGCGCCTCTT 400

RESULT 14
MUSURKI

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
source
CDS
misc_binding
polya_signal
BASE COUNT
ORIGIN

MUSURKI
Mus musculus uridine kinase mRNA, partial cds.
L31783
L31783.1 GI:471980
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1810)
Traut,T.W.
The functions and consensus motifs of nine types of peptide
segments that form different types of nucleotide-binding sites
Eur. J. Biochem. 222 (1), 9-19 (1994)
94259063
8200357
2 (bases 1 to 1810)
REFERENCE
Ropp,P.A. and Traut,T.W.
Cloning and expression of a cDNA encoding uridine kinase from mouse
brain
Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
97108719
8951040
Location/Qualifiers
1. .1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
<1. .783
/EC number="2.7.1.48"
/standard name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAB50568.1"
/db_xref="GI:471981"
/translation="RPPRPFLIGVSGTASGKSTVCKEIMELGNEVDRRQRKLV
LSODCYKVLTAOKAKALGOYNPHDPAPFNDLMHKTLMNIVSGKTVFPTDPT
HRLPFTTVYVADYVLPFGILVPTQERDMFHLRFPDTSVDRSLRVLADVRG
RDLQGLTQYIAVPAFEEELCLPTKATADVLIIPKVDNMVAINLIVQHIQDILNDL
CRNRGGPNRNMKRTFPPGDPGVLATGKSHLESSSRPH"
join(37. .63,412. .426,451. .465)
/bound_molecule="ATP at catalytic site"
1790 .1795
411 a 482 c 512 g 405 t

Query Match 2.2%; Score 36; DB 10; Length 1810;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 528 CCAGAGATCCGGACATGTTCCACCTGCGCCTCTT 563
Db 384 CCAGAGATCCGGACATGTTCCACCTGCGCCTCTT 419

RESULT 15
BC025146
LOCUS 1959 bp mRNA linear ROD 27-MAR-1997
DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
MGC:36231 IMAGE:4913412), complete cds.
BC025146
MGC.
BC025146.1 GI:19263563
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1959)
Strausberg,R.L., Feingold,E.A., Gronow,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Scapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Ueding, T. B., Tothiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loguigliano, N. A., Peters, G. J., Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, O. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Huiyk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butlerfield, J. S., Krzywinski, M. I., Skalska, U., Smalins, D. E., Scherch, A., Schein, J. E., Jones, S. J., and Marra, M. A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P. H., Garcia, A. M., Lu, X., Huiyk, S. W., Louised, H., Kowis, C. R., Sneed, A. J., Martin, R. G., Muzny, D. M., Navavati, A. N., Gibbs, R. A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAX Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES
SOURCE

gene
CDS

1. 1959
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_id="NCI_CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57. 890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translation="MASAGGSGSESAPADPPRPFLIGSGTASGSKTCEKIM ELTGNEVDRORLVLISODCFKVLTAOKAKLKGYNPDPAFNDLMHKTLM NIVEGTEVPTDFTVTRPTETTVYPAADVLFEGILVFOEIRDMFLALPVDV DSDVRLSRVLRVQGRGRDLEQITOTATYAKPAFEFCCLPTKIKIADVIIPRGVDNNV AINLIVOHIDILNGLCKRHGGPNGRNKHRTFPEPGHPVLATGKRSHLESSRP

BASE COUNT H
466 a 511 c 561 g 421 t
ORIGIN
Query Match 2.2%; Score 36; DB 10; Length 1959;
Best Local Similarity 100.0%; Pred. No. 6.4e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 528 CCAGAGATCCGGAGACATGTTCCACCTGCGCCTCTT 563
Db 491 CCAGAGATCCGGAGACATGTTCCACCTGCGCCTCTT 526

Search completed: November 25, 2003, 05:12:17
Job time : 6322.18 secs


```

Oy      61  CACCGGCCCTTCTGTATAGGGGTGAGCGGGGCACTGCGACAGGGGAAAGTGCACCGTGTGT 120
Db      155  CACGGGCCCTTCTGTATAGGGGTGAGCGGGGCACTGCGACAGGGGAAAGTGCACCGTGTGT 21.4
Oy      121  GAGAAAGTCATGAGAGTTGCTGTGGACAAAGAGGTGAAACAGCGGCACGGAGAGTGTCT 180
Db      215  GAGAAAGTCATGAGAGTTGCTGTGGACAAAGAGGTGAAACAGCGGCACGGAGAGTGTCT 27.4
Oy      181  ATCTGTAGCCAGAGACAGGTTCTTCAAGGTTCTGACGGGCAGACAGAAAGGCCAAGGCTTGG 240
Db      275  ATCTGTAGCCAGAGACAGGTTCTTCAAGGTTCTGACGGGCAGACAGAAAGGCCAAGGCTTGG 33.4
Oy      241  AAAAGACAGTACAAATTTTGGACCATCCAGATGCTTTGATAATGATTTGATGTCACAGGACT 300
Db      335  AAAAGACAGTACAAATTTTGGACCATCCAGATGCTTTGATAATGATTTGATGTCACAGGACT 33.4
Oy      301  CTGAAGAACATCGTGTGAGGGCAAAAACGAGTGAAGGTGCCGACCTATGATTTTGTGACAC 360
Db      395  CTGAAGAACATCGTGTGAGGGCAAAAACGAGTGAAGGTGCCGACCTATGATTTTGTGACAC 45.4
Oy      361  TCAAGGTTACCAAGACCAACGGTGGTCTACCTGGCGGACGTGGTCTGTGTTGAGGGCATC 420
Db      455  TCAAGGTTACCAAGACCAACGGTGGTCTACCTGGCGGACGTGGTCTGTGTTGAGGGCATC 51.4
Oy      421  TTGGTGTCTTACAGCCAGAGAGATCCGGGACATGTTTCAACTGGCGGCTTTGTGTGAACAC 480
Db      515  TTGGTGTCTTACAGCCAGAGAGATCCGGGACATGTTTCAACTGGCGGCTTTGTGTGAACAC 57.4
Oy      481  GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCCCGGAGGAGGACCTG 540
Db      575  GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCCCGGAGGAGGACCTG 63.4
Oy      541  GAGCAGATTTTGACCGACGTACACACCTTCTGTGAAGCCGGCTTTGAGGAAGTTCTGCTTG 600
Db      635  GAGCAGATTTTGACCGACGTACACACCTTCTGTGAAGCCGGCTTTGAGGAAGTTCTGCTTG 69.4
Oy      601  CCGACAAAGAAAGTATGCCGATGATCATCCACAGAGAGTGGACAATATGTTGGCATC 660
Db      695  CCGACAAAGAAAGTATGCCGATGATCATCCACAGAGAGTGGACAATATGTTGGCATC 75.4
Oy      661  AACCTGATCGTGCAGACATCCAGACATTTGTAATGTGTACATCTGCATAATGACACCGA 720
Db      755  AACCTGATCGTGCAGACATCCAGACATTTGTAATGTGTACATCTGCATAATGACACCGA 81.4
Oy      721  GAGAGGTCCAATGAGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGGAGCAACCTCTGG 780
Db      815  GAGAGGTCCAATGAGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGGAGCAACCTCTGG 87.4
Oy      781  ATGTGTACCTCTGCGCAAAACGGTTCACATTTGGAGTCCAGACAGACCCCACTGA 834
Db      875  ATGTGTACCTCTGCGCAAAACGGTTCACATTTGGAGTCCAGACAGACCCCACTGA 928

RESULT 2
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen >
; APPLICANT: Johnson, Randall
; TITLE OR INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: G950020
; CURRENT APPLICATION NUMBER: US/09/536,647A
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)

```

```

US-09-536-647-1
Query Match 99.8%; Score 832.4; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 6.4e-205;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0

1 ATGCGCTTCGCGGGAGGCGGAGAGCTGCGAGAGCCCGCGCGGAGGCGGAGCGTCCGAC 60
1 ATGCGCTTCGCGGGAGGCGGAGAGCTGCGAGAGCCCGCGCGGAGGCGGAGCGTCCGAC 60
61 CAGCGGCCCTTCTTGATAGGGGTGAGCGCGCGCACTGCGAGCGGAACTCGACCGTGT 120
61 CAGCGGCCCTTCTTGATAGGGGTGAGCGCGCGCACTGCGAGCGGAACTCGACCGTGT 120
121 GAGAAAGATCATGAGAGTCTGTGGGACAGAACGAGGTGGAACAGCGGCAACGGAAAGTGTCTC 180
121 GAGAAAGATCATGAGAGTCTGTGGGACAGAACGAGGTGGAACAGCGGCAACGGAAAGTGTCTC 180
181 ATCTGAGCCAGAGCAGGTTCTTACAAGTCTTGACGCGAGAGCAGAAAGCCAGGCTTG 240
181 ATCTGAGCCAGAGCAGGTTCTTACAAGTCTTGACGCGAGAGCAGAAAGCCAGGCTTG 240
241 AAAGGACAGTACAAATTTTGACATCCAGATGCTTTGTAATGATTGATGCAAGAGCT 300
241 AAAGGACAGTACAAATTTTGACATCCAGATGCTTTGTAATGATTGATGCAAGAGCT 300
301 CTGAAGAAACATCTGTGAGAGGCGGAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
301 CTGAAGAAACATCTGTGAGAGGCGGAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
361 TCAAGGTTACAGAGACCCAGCGGTGTCAACCGTGGGAGCTGTGTTCTTTGAGGGCATC 420
361 TCAAGGTTACAGAGACCCAGCGGTGTCTCAACCGTGGGAGCTGTGTTCTTTGAGGGCATC 420
421 TTGTTGTTCTTACAGCGAGAGATCCGGGACATGTTCCACTGCGCTTTGCTGTGACAC 480
421 TTGTTGTTCTTACAGCGAGAGATCCGGGACATGTTCCACTGCGCTTTGCTGTGACAC 480
481 GACTCCGACGTCAGGCTGTCTTCAAGAGTTCTCCGGGACGTCGCGCGGAGGGAGGACCTG 540
481 GACTCCGACGTCAGGCTGTCTTCAAGAGTTCTCCGGGACGTCGCGCGGAGGGAGGACCTG 540
541 GAGCAGATTCGAGCGAGTACACCACTTCGTGAAGCCGGGCTTCGAGAGTTCGCGCTG 600
541 GAGCAGATTCGAGCGAGTACACCACTTCGTGAAGCCGGGCTTCGAGAGTTCGCGCTG 600
601 CCGACAAAGAGTATGCGGATGTATCATCCACGAGGAGTGAACAATATGTTGCCATC 660
601 CCGACAAAGAGTATGCGGATGTATCATCCACGAGGAGTGAACAATATGTTGCCATC 660
661 AACCTGATCTGCGAGCACATCCAGACATTTCTGAATGTTGACATCTGCAATATGCGACGA 720
661 AACCTGATCTGCGAGCACATCCAGACATTTCTGAATGTTGACATCTGCAATATGCGACGA 720
721 GGAGGTCCTAATGGCGGAGCTTACAAGGGGACCTTTCTGAGCCAGGGGACCAACCTGAG 780
721 GGAGGTCCTAATGGCGGAGCTTACAAGGGGACCTTTCTGAGCCAGGGGACCAACCTGAG 780
781 ATGCTGACCTTGGCGAAACGGTTCACATTTGAGAGTCCAGACAGACCCCACTGA 834
781 ATGCTGACCTTGGCGAAACGGTTCACATTTGAGAGTCCAGACAGACCCCACTGA 834

RESULT 3
US-09-221-017B-368
; Sequence 368, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Roser, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & POERSTER
;

```



```

STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221.017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PPI182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PPI546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA: PPI2911
FILING DATE: 09-APR-1998
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: MONROY, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 5687 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...5687
US-09-221-017B-368

Query Match      12.8%; Score 106.8; DB 4; Length 5687;
Best Local Similarity 54.3%; Pred. No. 4.1e-18;
Matches 238; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

```

```

QY 545 AGATTCGACGAGTACACACCTTGTAAGCCGCGCTTCGAGGAGTTCTGCGCCGA 604
DB 5154 CGGTATTGGACATATCTTCAAGCGTTCCGCTTCGACGAGGATTTGTGAAACCAT 5213
QY 605 CAAAGAAGTATGCCGATGATCATCCACGAGAGTGAACATATGTTGCCATCAACC 664
DB 5214 CCAAGGCGTACGCCATCTGATCATTCGGAAGGTGATTCAATTGGGCGCTCTCAC 5273
QY 665 TGATCGTGAACACATCC 682
DB 5274 TCCTGTCCAAAAATCC 5291

RESULT 4
US-08-961-527-76/c
Sequence 76, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 10011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-76

Query Match      12.2%; Score 101.8; DB 4; Length 10011;
Best Local Similarity 53.2%; Pred. No. 9.3e-17;
Matches 239; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

```

```
Db      9821 GGACACAGCTCTGCGCATTTTGATGATATCAATTTTGTGGATTCGAGTACGATGT 9762
Qy      492 CAGGCTGTCTCGAAGAGTTCTCCGGGA---CGTGGCCGAGGAGGAGGAGCTTGAGCAGAT 548
Db      9761 GCGCATTTATTCGTGATCAAGCGTGATGATGAGAGCGTGGCGGTGACCTTGATAGCGT 9702
Qy      549 TCTGACCGCATGACACACCTTCGTGAAGCCGGCTTGAGAGAGTTCTGCTGCCGAAAAA 608
Db      9701 TATTAAACAGTACTTAGGTGTGTCAACCAATGTACACCGATTATCGAGTCAACTAA 9642
Qy      609 GAATGTGCGCATGTGATGATCCACGAGAGATGAGCAATATGTTCCATCACTGAT 668
Db      9641 GCGTATGCTGATATCTGATCTCTGAAGGGGTTAGCAATACGTGGCTATGACTGTT 9582
Qy      669 CGTGACGACATCCAGACATTTGGAATG 697
Db      9581 GACGACCAAGATTCGAAAGATTTTGAAG 9553

RESULT 5
US-09-134-001C-781
; Sequence 781, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 781
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-781

Query Match      10.3%; Score 86.2; DB 4; Length 651;
Best Local Similarity 51.5%; Pred. No. 4,1e-13;
Matches 224; Conservative 0; Mismatches 208; Indels 3; Gaps 1;

Qy      253 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAGAGATC 312
Db      196 AATTATGACCACTTGGATTTGATATGATTTACTATTATTAATTTAAAGATTTA 255
Qy      313 GTGAGGGGCAAAACGGTGGAGGTGCCGACTATATTTTGTGACACACTCAAGTTACA 372
Db      256 AGAATATGAAAAACGATAGAGTACTAGTACGATTATCTCAACATACACGATGATAA 315
Qy      373 GAGACACGGGTGCTACCTCCGCGAGAGTGTCTGTTGAGGGGATCTGGTGTCTAC 432
Db      316 GAAACAATTGATTTGATTCAAAAGATGTATATCGAGAAGTATCTTTTGTTGAA 375
Qy      433 AGCCAGAGATCCGAGCATGTTCCACTGCGCTCTTCTGTGACACCGACTCCGAGCTC 492
Db      376 AACAACAACATTACGAGCATGATGATGTGAATTTTATGTCATACGATGATGATTTA 435
Qy      493 AGGCTGTCTGAAGAGTTCTCCGGGA---CGTGGCCGAGGAGGAGACTGAGACAGATT 549
Db      436 CGAATTTTACGTAGGCTTACAAGAGATCTAAGAAGCGTGTCTGTAACAATGGAATCAGTA 495
Qy      550 CTGACGAGTACACACCTTGTAAGCGGCGCTTTCAGAGAGTCTGCTGCCGACCAAG 609
Db      496 ATTATATCAATATCTTAATGTAGTAAGACCTATGATGAGCAATTTATTGAACTTACAAA 555
Qy      610 AAGTATGCGCATGTGATCATCCACGAGAGAGTGAACAATATGTTGCATCAACCTGATC 669
Db      556 AAGCATGTGATCATATATCTCTGAAGAGAGCAATAAGTTGCAATGATATATATG 615
```

```
Qy      670 GTGACACATGACG 684
Db      616 ACTATAAATCCAG 630

RESULT 6
US-09-107-532A-1370
; Sequence 1370, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCIT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-5007
; INFORMATION FOR SEQ ID NO: 1370:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (b) LOCATION 1...633
; SEQUENCE DESCRIPTION: SEQ ID NO: 1370:
US-09-107-532A-1370

Query Match      8.8%; Score 73.4; DB 4; Length 633;
Best Local Similarity 50.1%; Pred. No. 7.9e-10;
Matches 210; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

Qy      253 AATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAGAGATC 312
Db      178 AATTAGATCACTCATTTGCTTTGATATGATCTCTGATTCAGCATGAGCGGACTTA 237
Qy      313 GTGAGGGGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACACTCAAGTTACA 372
Db      238 CTGAACATATAAGCCATTGAAAAACCGTGTATGACTATGTACCGCATACAAAGTCAAG 297
Qy      373 GAGACACGGGTGTACCTCTGGGAGCGTGTCTGTTGAGGGGCACTTGTGTCTAC 432
```

Db 298 GCAACGATCATTCAGAACCAAAAAGATGATCATTTAGAGGAATCTGATTTAGAA 357
Qy 433 AGCCAGAGATCCGGGACATGTTCCACCTCGCTCTTGAGACACCGACCTCCGAGTC 492
Db 358 GACGAACGCTGCGTGGCTTGAATGAGATATAAAGTATACGTATGATGATGATGAT 417
Qy 493 AGGCTGCTCGAAGAGTTCTCCGGACGTG---CGCCGAGGAGGAGCTGAGCAGATT 549
Db 418 CGTATCATTTGCTGTATCAACCGGATATGAGAGACGGAGCAGCATTTGACTCCGTG 477
Qy 550 CTGACGAGTACACCACTTCGTGAAGCCGCGCTTCGAGAGATTCTGCTCCGACAAAG 609
Db 478 ATCGAACATATTTGACAGTTGTTAAACGATGATCATCAATTTATCGAACCAAG 537
Qy 610 AAGTATGCCATGTGATCATCCACGAGAGTGAGATATGTTGCTGATCAACCTGAT 668
Db 538 CGCTATGCGGATATCATTTGTTCCAGAGGCGGAGAAAACACGTTGCTATTTGAT 596

RESULT 7
US-09-198-452A-1/c
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifffals, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (15001)-(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (30001)-(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (45001)-(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (60001)-(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (75001)-(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (90001)-(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (105001)-(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (120001)-(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (135001)-(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (150001)-(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (165001)-(180000)
OTHER INFORMATION: n=a or c or g or t

NAME/KEY: misc feature
LOCATION: (180001)-(195000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (195001)-(210000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (210001)-(225000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (225001)-(240000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (240001)-(255000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (255001)-(270000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (270001)-(285000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (285001)-(300000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (300001)-(315000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (315001)-(330000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (330001)-(345000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (345001)-(360000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (360001)-(375000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (375001)-(390000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (390001)-(405000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (405001)-(420000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (420001)-(435000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (435001)-(450000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (450001)-(465000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (465001)-(480000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (480001)-(495000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (495001)-(510000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (510001)-(525000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature
LOCATION: (525001)-(540000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc feature

LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (690001)..(705000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (705001)..(720000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (720001)..(735000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (735001)..(750000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (750001)..(765000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (765001)..(780000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (780001)..(795000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (795001)..(810000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (810001)..(825000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (825001)..(840000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (840001)..(855000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (855001)..(870000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (870001)..(885000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (885001)..(900000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (900001)..(915000)

OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
Query Match 6.7%; Score 55.8; DB 4; Length 1230025;
Best Local Similarity 49.9%; Pred. No. 0.00028;
Matches 197; Conservative 0; Mismatches 192; Indels 6; Gaps 2;
Qy 248 AGTACATTTTTCACATCCGATGCTTGATATGATTTGATGACAGACTCTGAAGA 307
Db 835626 ATTTAATTTGGATCATTCCGAGCCCTTGATATGATTTGATGACATTAAC 835567
Qy 308 ACATCGTGAGGCGCAAAACGGTGGAGTCCGACCTATGATTTTGACA---CACTCAA 364
Db 835566 GTCTAAATAATATGATGATGTCGACGCCCAATTTTATTTTATGTAATGAT 835507
Qy 365 GGTTCACGAGACCAAGGTGCTTACCTCGGACGTGTTGTTGAGGCGATCTTG 424
Db 835506 CTAAACGGAGATGAAACGATCTATCCATCTAAAGTATTTCTTTGTAAGGATTTCTGG 835447
Qy 425 TGTTCATACGCCGAGATCCGGACATGTTCACTCGGCGCTTCTGTGACACCGACT 484
Db 835446 TCTTTGAAATCAAGAACTTAGATCTTAGATTTAGATCTTTGTAACACCGATG 835387
Qy 485 CCGACGTCAAGCTGTCTCGAAGATTCTCCGGACGTGC---GCCGAGGAGGAGACTTG 541
Db 835386 CTGATGAAGAGATACACGCCGATGTTGAGATGTTCAAGAACAGAGATAGCTG 835327
Qy 542 AGCAGATTCTGACGACGATACCACTTGCTGMAAGCGGCTTCGAGAGTTCTGCCTGC 601
Db 835326 ACTGCATCATGCTCGTTATCTTTCTTAGTGTAAGCCTTAGCATGAGAAATTTATAGAG 835267
Qy 602 CGACAAAGAGTATGCCATGATGATCATCCACGA 636
Db 835266 CGACTCGAATATGCTGATATCATTTGATACATGA 835232
RESULT 8
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:


```
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 4092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 200...3028
OTHER INFORMATION:
NAME/KEY: Human ROR2
LOCATION: 1...4092
OTHER INFORMATION:
US-08-469-537A-106

Query Match      6.0%; Score 50.4; DB 2; Length 4092;
Best Local Similarity 48.9%; Pred. No. 0.0012;
Matches 135; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 129 CATGAGTTGCTGGACAGAAAGAGGTGAAACAGCGGCGAGCGAAGGTGCTCTGAG 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1552 CATGAAATGCCCTCATTTAACAGACACAAAGCCAACTCAAAAGATCAAGCTGTC 1611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 189 CCAGACAGGTTCTTCAAGGCTCTGACGGCAGACAGAGCCCAAGGCTTGAAGAACA 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1612 TGGCGTAGGTTCTATGAGAGGCTGGAGAGGACCGGTTTGGAAAGTCTACAAAGGTCA 1671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 249 GTACAAATTTTGACCATCCAGATGCTCTTGTATATGATTTGACACAGACTCTGAAGAA 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1672 CCTGTCGGCCCTGCCCGGGGGAGAGACCCAGAGCTGGCCATCAAAAGCTGAAGGA 1731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 CATGTGAGGCGAAACGCTGAGAGTGCACACTATGATTTTGTACACACTCAAGGTT 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1732 CAAGCGAGAGGGCCCTGCGGAGGAGTTCCGGCATGAGGCTATGCTGGAGCAGCGCT 1791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 ACCAGAGCACCGGNGTCTACCTCGGCGAGCGTGT 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1792 GCAACACCCCAAGTCTGCTGCTGCGGCGTGT 1827
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

```
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match      5.5%; Score 46; DB 1; Length 7218;
Best Local Similarity 4.7%; Pred. No. 0.019;
Matches 19; Conservative 216; Mismatches 171; Indels 0; Gaps 0;

QY 71 TCCTGATAGGGGTGACGGCGGCACTGCCAGCGGGAAGTCGACCGTGTGAGAAAGATCA 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 TGAAGTTCTGGACAGAAAGAGGTGAAACAGCGGCGAGGAAAGTGTCTGAGCC 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 AGGACAGTCTTCAAGGCTCTGACGGCAGAGAGAGCAAGGCTTGAAGAAGCACT 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 ACAATTTGACCATCCAGATGCTCTTGTATATGATTTGACACAGACTCTGAAGAACA 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 TCGTGAAGGCGAAACGCTGAGAGTGCACACTATGATTTTGTACACACTCAAGTTAC 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 371 CAGAGACACAGGCGTCACTCGCGGAGCGTGTCTGTTAGGGCATCTGTTCT 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 431 ACAGCCAGAGATCCGGAATGTTCACTGCGCTTCTGTGA 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 1077 RRRRRRRRRATCGCAAGCTCCCTGACCTCGACGCAACTGCGA 1032

RESULT 12
US-09-252-991A-931
; Sequence 931, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 931
; LENGTH: 1437
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-931

Query Match 5.2%; Score 43; DB 4; Length 1437;
Best Local Similarity 52.5%; Pred. No. 0.067;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGAGACCGACTCGGACG 490
Db 218 ACAGCTGCTGTTCCGACCACTAGCACTGCTGCTACCGACTTGGGACCGCGCAAG 277
Qy 491 TCAGCTGTCTCGAAGATTCTCGGAGCGTGGCCGAGGAGGAGCCTGAGCAGATTTC 550
Db 278 GCTGGCGCAACTCCGGGACGTGGCGGACGAACTCAAGCCGCAAGGGGATCGAGCTGTGG 337
Qy 551 TGACGAGTACACCACTTGTGTGAAGCGGCTTTCAGAGATTCTGCTCCGCAAG 609
Db 338 TGTCTACCAAGCGGACCGCGGCTGTGTGAACCGCGAAGCTCAGCCCGGAAAG 396

RESULT 13
US-09-252-991A-1020
; Sequence 1020, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1020
; LENGTH: 1647
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1020

Query Match 5.2%; Score 43; DB 4; Length 1647;
Best Local Similarity 52.5%; Pred. No. 0.07;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGAGACCGACTCGGACG 490
Db 720 ACAGCTGCTGTTCCGACCACTAGCACTGCTGCTACCGACTTGGGACCGCGCAAG 779

Qy 491 TCAGCTGTCTCGAAGATTCTCCGGGACGTGGCGGAGGAGGACCTGAGCAGATTTC 550
Db 780 GCTGGCGCAACTCCGGGACCTGCGGACGAACTCAAGGCGCAAGGGATCGAGCTGTGG 839
Qy 551 TGACGAGTACACCACTTGTGTGAAGCGGCTTTCAGAGATTCTGCTCCGCAAG 609
Db 840 TGTCTACCAAGCGGACCGCGGCTGTGTGAACCGGAGAACTCAGCCCGGAAAG 898

RESULT 14
US-09-252-991A-1036/c
; Sequence 1036, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1036
; LENGTH: 2847
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1036

Query Match 5.2%; Score 43; DB 4; Length 2847;
Best Local Similarity 52.5%; Pred. No. 0.083;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 431 ACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGAGACCGACTCGGACG 490
Db 914 ACAGCTGCTGTTCCGACCACTAGCACTGCTGCTACCGACTTGGGACCGCGCAAG 855
Qy 491 TCAGCTGTCTCGAAGATTCTCCGGAGCGTGGCCGAGGAGGAGCCTGAGCAGATTTC 550
Db 854 GCTGGCGCAACTCCGGGACGTGGCGGACGAACTCAAGCCGCAAGGGGATCGAGCTGTGG 795
Qy 551 TGACGAGTACACCACTTGTGTGAAGCGGCTTTCAGAGATTCTGCTCCGCAAG 609
Db 794 TGTCTACCAAGCGGACCGCGGCTGTGTGAACCGCGAAGCTCAGCCCGGAAAG 736

RESULT 15
US-09-252-991A-12919
; Sequence 12919, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12919
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12919

Query Match 4.8%; Score 40.4; DB 4; Length 444;
Best Local Similarity 53.1%; Pred. No. 0.22;
Matches 86; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 10 GCGGAGGCGAAGACTGCGAGAGCCCCCGCGCCCGAGAGCCGACCGTCCGACACCGAGGCCC 69
 Db 110 GTGAAAGACGATGATCCCGGAATTCCCGCCCGACCGCGAGCGGCTTCCAGCGCTTC 169
 OY 70 TTCTGATAGGAGGTGAGCGCGGCACTGCCAGCGGAAATCGACCGTGTGAGAGATC 129
 Db 170 TACCGGCGCGGCGACAGCCCCCGGCGAGCGGCTGGGGCTGGCGATGTCGCGAGGTCTGC 229
 OY 130 ATGAGTTGTCGGGACAGAACGAGGTGGAACAGCGCGACGCG 171
 Db 230 CGGGGCGACCGGCGCGAGATCCAGCTGACCAAGGCGAACTG 271

Search completed: November 25, 2003, 02:04:04
 Job time : 69.681 secs

Db 95 ATGGCTTC

95 AIGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCCCGCGCGAGGCGGACCGTCCGCAC 154

```
QY 61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAGTGCACCGTGTGT 120
DB 155 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAGTGCACCGTGTGT 214
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACGGGAAGTGTGTC 180
DB 215 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACGGGAAGTGTGTC 274
QY 181 ATCTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGAGGACAGAAAGGCCAGCTTG 240
DB 275 ATCTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGAGGACAGAAAGGCCAGCTTG 334
QY 241 AAGGACAGTACATTTTGAACATCCAGATGCTCTTGAATATGATTTGATGACAGACT 300
DB 335 AAGGACAGTACATTTTGAACATCCAGATGCTCTTGAATATGATTTGATGACAGACT 394
QY 301 CTGAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 360
DB 395 CTGAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 454
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTCGCGACGTTGTTCTGTTGAGGCAATC 420
DB 455 TCAAGGTTACAGAGACCAAGTGTCTACCTCGCGACGTTGTTCTGTTGAGGCAATC 514
QY 421 TTGGGTCTTACAGCGAGGAGATCCGGGACATGTTCCACCTGGGCTCTTCTGTGACACC 480
DB 515 TTGGGTCTTACAGCGAGGAGATCCGGGACATGTTCCACCTGGGCTCTTCTGTGACACC 574
QY 481 GACTCCGACCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGACCTG 540
DB 575 GACTCCGACCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGACCTG 634
QY 541 GAGCAGATTCTGACGCGATACCACTTCTGTGAAGCCGCTTCTGAGAGTTTGTGCTG 600
DB 635 GAGCAGATTCTGACGCGATACCACTTCTGTGAAGCCGCTTCTGAGAGTTTGTGCTG 694
QY 601 CCAGCAAGAAGTATGCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 695 CCAGCAAGAAGTATGCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 754
QY 661 AACCTGATCTGACGACATCCAGAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 720
DB 755 AACCTGATCTGACGACATCCAGAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 814
QY 721 GAGAGGTTCAATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCGAGGAGACCTTGGG 780
DB 815 GAGAGGTTCAATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCGAGGAGACCTTGGG 874
QY 781 ATGCTGACCTCTGGCAACGCTCATTTTGAAGTCCAGACAGACCCCACTGA 834
DB 875 ATGCTGACCTCTGGCAACGCTCATTTTGAAGTCCAGACAGACCCCACTGA 928

RESULT 2
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: G950020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(780)
```

```
US-09-536-647-1
Query Match 93.9%; Score 783; DB 4; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTTCGGCGGAGGCGGAAGCATGCGAGAGCCCGCGCGAGGCGGACCGCTCCGAC 60
DB 1 ATGAGCTTCGGCGGAGGCGGAAGCATGCGAGAGCCCGCGCGAGGCGGACCGCTCCGAC 60
QY 61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAGTGCACCGTGTGT 120
DB 61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAGTGCACCGTGTGT 120
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACGGGAAGTGTGTC 180
DB 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACGGGAAGTGTGTC 180
QY 181 ATCTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGAGGACAGAAAGGCCAGCTTG 240
DB 181 ATCTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGAGGACAGAAAGGCCAGCTTG 240
QY 241 AAGGACAGTACATTTTGAACATCCAGATGCTCTTGAATATGATTTGATGACAGACT 300
DB 241 AAGGACAGTACATTTTGAACATCCAGATGCTCTTGAATATGATTTGATGACAGACT 300
QY 301 CTGAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 360
DB 301 CTGAGAACATCTGAGAGGCGAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 360
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTCGCGACGTTGTTCTGTTGAGGCAATC 420
DB 361 TCAAGGTTACAGAGACCAAGTGTCTACCTCGCGACGTTGTTCTGTTGAGGCAATC 420
QY 421 TTGGGTCTTACAGCGAGGAGATCCGGGACATGTTCCACCTGGGCTCTTCTGTGACACC 480
DB 421 TTGGGTCTTACAGCGAGGAGATCCGGGACATGTTCCACCTGGGCTCTTCTGTGACACC 480
QY 481 GACTCCGACCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGACCTG 540
DB 481 GACTCCGACCTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGACCTG 540
QY 541 GAGCAGATTCTGACGCGATACCACTTCTGTGAAGCCGCTTCTGAGAGTTTGTGCTG 600
DB 541 GAGCAGATTCTGACGCGATACCACTTCTGTGAAGCCGCTTCTGAGAGTTTGTGCTG 600
QY 601 CCAGCAAGAAGTATGCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 601 CCAGCAAGAAGTATGCGATGTGATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATCTGACGACATCCAGAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 720
DB 661 AACCTGATCTGACGACATCCAGAGACATTTCTGAATGTGACATCTGCAATGTGACCGA 720
QY 721 GAGAGGTTCAATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCGAGGAGACCTTGGG 780
DB 721 GAGAGGTTCAATGGGCGGAGCTTCAAGCGGACCTTTTCTGAGCCGAGGAGACCTTGGG 780
QY 781 ATGCTGACCTCTGGCAACGCTCATTTTGAAGTCCAGACAGACCCCACTGA 834
DB 781 ATGCTGACCTCTGGCAACGCTCATTTTGAAGTCCAGACAGACCCCACTGA 834

RESULT 3
US-08-098-141-1
; Sequence 1, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
```

STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02119
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617:227-5020
TELEFAX: 617:227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

Query Match 2.5%; Score 21; DB 1; Length 3060;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 TCATCCTGAGCCAGACAGGT 199
Db 1344 TCATCCTGAGCCAGACAGGT 1364

RESULT 4
US-09-252-991A-4602/c
Sequence 4602, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 4602
LENGTH: 600
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4602

Query Match 2.0%; Score 17; DB 4; Length 600;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 TCATGAGTGTGCGGGG 144
Db 251 TCATGAGTGTGCGGGG 235

RESULT 5

US-09-276-531-4
Sequence 4, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 887 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGPUT03
CLONE: 1472268
US-09-276-531-4

Query Match 2.0%; Score 17; DB 3; Length 887;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 GTGGTCTGTGTGAGGG 416
Db 423 GTGGTCTGTGTGAGGG 439

RESULT 6
US-09-134-001C-2296
Sequence 2296, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964

;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 2296
;; LENGTH: 1329
;; TYPE: DNA
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2296

Query Match 2.0%; Score 17; DB 4; Length 1329;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 ATTTGACCATCCAGAT 270
 |||||
DB 530 ATTTGACCATCCAGAT 546

RESULT 7
US-09-801-861-1
; Sequence 1, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-1

Query Match 2.0%; Score 17; DB 4; Length 2203;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 782 TGCTGACCTCTGGCAA 798
 |||||
DB 25 TGCTGACCTCTGGCAA 41

RESULT 8
US-09-252-991A-4346
; Sequence 4346, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4346
; LENGTH: 2493
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4346

Query Match 2.0%; Score 17; DB 4; Length 2493;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 TCATGAGTTGCTGGGA 144
 |||||
DB 2293 TCATGAGTTGCTGGGA 2309

RESULT 9
US-09-962-665-4
; Sequence 4, Application US/09962665
; Patent No. 6537759
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: FOLYPOLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 16
; SEQ ID NO 4
; LENGTH: 2816
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 175..1067
; OTHER INFORMATION: n = g or a
; NAME/KEY: misc_feature
; LOCATION: 341
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 791..1997, 2618, 2653
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1337
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 2107
; OTHER INFORMATION: nucleotide in position 2107 is g, or absent
; NAME/KEY: misc_feature
; LOCATION: 2583
; OTHER INFORMATION: n = t or g
US-09-962-665-4

Query Match 2.0%; Score 17; DB 4; Length 2816;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 GCGGCACTGCCACGGG 105
 |||||
DB 1460 GCGGCACTGCCACGGG 1476

RESULT 10
US-09-045-201A-1
; Sequence 1, Application US/09045201A
; Patent No. 6110718
; GENERAL INFORMATION:
; APPLICANT: Shisheva, Assia
; TITLE OF INVENTION: No. 6110718el Mammalian Putative
; TITLE OF INVENTION: Phosphatidylinositol-4-Phosphate-5-Kinase
; NUMBER OF SEQUENCES: 2

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
;; STREET: P.O. Box 828
;; CITY: Bloomfield Hills
;; STATE: Michigan
;; COUNTRY: U.S.A.
;; ZIP: 48303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/045,201A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, Dean F
;; REFERENCE/DOCKET NUMBER: 4981-098431
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (248) 641-1600
;; TELEFAX: (248) 641-0270
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6297 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-045-201A-1

Query Match 2.0%; Score 17; DB 3; Length 6297;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GCGCGAGGCGGAGACT 25
Db 116 GCGCGAGGCGGAGACT 132

RESULT 11
US-09-619-062-1
; Sequence 1, Application US/09619062
; Patent No. 6406875
; GENERAL INFORMATION:
; APPLICANT: Shisheva, Assia
; TITLE OF INVENTION: No. 6406875e1 Mammalian Putative
; TITLE OF INVENTION: Phosphatidylinositol-4-Phosphate-5-Kinase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/619,062
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/045,201
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Dean F
; REFERENCE/DOCKET NUMBER: 4981-098431
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 641-1600
; TELEFAX: (248) 641-0270

;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6297 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-09-619-062-1

Query Match 2.0%; Score 17; DB 4; Length 6297;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GCGCGAGGCGGAGACT 25
Db 116 GCGCGAGGCGGAGACT 132

RESULT 12
US-08-311-731A-132/C
; Sequence 132, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 132:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36412 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LAPRAE
; US-08-311-731A-132

Query Match 2.0%; Score 17; DB 4; Length 36412;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GCGCGACTGCCAGCGG 104
Db 25517 GCGCGACTGCCAGCGG 25501

```
RESULT 13
US-08-311-731A-130/C
; Sequence 130, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36941 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
; US-08-311-731A-130

Query Match          2.0%; Score 17; DB 4; Length 36941;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      219 AGACGAGAGGCCAAG 235
DB      13109 AGACGAGAGGCCAAG 13093

RESULT 14
US-09-918-686-2
; Sequence 2, Application US/09918686
; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```

```
LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match          2.0%; Score 17; DB 4; Length 51719;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      290 TGCACAGACTCTGGAAG 306
DB      35965 TGCACAGACTCTGGAAG 35981

RESULT 15
US-09-801-861-3
; Sequence 3, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
; US-09-801-861-3

Query Match          2.0%; Score 17; DB 4; Length 53332;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      782 TGCTGACCTCTGGCAAA 798
DB      2025 TGCTGACCTCTGGCAAA 2041

Search completed: November 25, 2003, 06:54:23
Job time : 68.3596 secs
```


FEATURES
Source Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 2.1e-172;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGCGAAGACTCGAGAGCCCCCGCGCGGACCGCCGAC 60
Db 1 ATGGCTTCGGGGGAGGCGAAGACTCGAGAGCCCCCGCGCGGACCGCCGAC 60
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCGACGCGGAAGTCGACCGTGT 120
Db 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCGACGCGGAAGTCGACCGTGT 120
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 180
Db 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 180
QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGGACGAGAGCCAAAGCCTTG 240
Db 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGGACGAGAGCCAAAGCCTTG 240
QY 241 AAAGGACAGTACATTTTGAACATCCAGATCCATCTTTGATTAATGATTGATCAGAGACT 300
Db 241 AAAGGACAGTACATTTTGAACATCCAGATCCATCTTTGATTAATGATTGATCAGAGACT 300
QY 301 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGAGTCCGACCTATGATTTTGTGACACAC 360
Db 301 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGAGTCCGACCTATGATTTTGTGACACAC 360
QY 361 TCAAGGTTTACAGAGACCAAGGTGTCTACCCCTCGGACGCGGTTCTGTTGAGGGATC 420
Db 361 TCAAGGTTTACAGAGACCAAGGTGTCTACCCCTCGGACGCGGTTCTGTTGAGGGATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCTCTTCTGTGACACC 480
Db 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCTCTTCTGTGACACC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGGACGTCGCGGACGAGGGAGGACCTG 540
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGGACGTCGCGGACGAGGGAGGACCTG 540
QY 541 GAGCAGATTCTGAGGACGATACACACCTTCGTAAGCGCGCTTCGAGGAGTTCTGCTG 600
Db 541 GAGCAGATTCTGAGGACGATACACACCTTCGTAAGCGCGCTTCGAGGAGTTCTGCTG 600
QY 601 CCGCAAAAGAGTATGCGATGTATATCCACGAGAGTGAACAATATGTTGCCATC 660
Db 601 CCGCAAAAGAGTATGCGATGTATATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATCTGACGACATCTCAGACATTTGAAATGATGATCTGCAATATGACACGA 720
Db 661 AACCTGATCTGACGACATCTCAGACATTTGAAATGATGATCTGCAATATGACACGA 720
QY 721 GGAAGGTTCATAGGCGGAGCTACAGAGGACCTTTTCTGAGCCAGGGGACCAACCTGGG 780
Db 721 GGAAGGTTCATAGGCGGAGCTACAGAGGACCTTTTCTGAGCCAGGGGACCAACCTGGG 780
QY 781 ATGCTGACCTCTGGCAAAAGGTCATATTTGAGTTCAGACAGACCAACCCACTGA 834
Db 781 ATGCTGACCTCTGGCAAAAGGTCATATTTGAGTTCAGACAGACCAACCCACTGA 834

RESULT 2
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002

DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
Glucksmann, M.A.
57658, a human uridine kinase and uses thereof
Patent: WO 0202761-A 1 10-JAN-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)

FEATURES
Source
1. .1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGEDCESEPAEPADRPHPORFLIGVSGTASGKSTCEKIM
ELIGONEBOROKVILISODREFYKVLTAEOAKALKGOYNDPDPAPFMDLHRTLK
NIVEGTVRVPYDPTVHSRLPETTVYVADVLVPSGLIVFSQIRDMFHLRIYDT
DSVRLSRVLADYRGRDLBOILTYTTTFFVAPAEFELPTKTIADVILIPRGVDMV
ALNLTVOHLDILINDICMHRGSGNSGRYKRTFSBPDHPALTSGRSHLESSRP
H"

BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGCGAAGACTCGAGAGCCCCCGCGCGGACCGCCGAC 60
Db 94 ATGGCTTCGGGGGAGGCGAAGACTCGAGAGCCCCCGCGCGGACCGCCGAC 153
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCGACGCGGAAGTCGACCGTGT 120
Db 154 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTCGACGCGGAAGTCGACCGTGT 213
QY 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 180
Db 121 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 180
QY 214 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 273
Db 214 GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 273
QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGGACGAGAGCCAAAGCCTTG 240
Db 274 ATCTGAGCCAGGACAGGTTCTACAGAGTCTCGACGCGGACGAGAGCCAAAGCCTTG 333
QY 241 AAAGGACAGTACATTTTGAACATCCAGATCCATCTTTGATTAATGATTGATCAGAGACT 300
Db 334 AAAGGACAGTACATTTTGAACATCCAGATCCATCTTTGATTAATGATTGATCAGAGACT 393
QY 301 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGAGTCCGACCTATGATTTTGTGACACAC 360
Db 394 CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGAGTCCGACCTATGATTTTGTGACACAC 453
QY 361 TCAAGGTTTACAGAGACCAAGGTGTCTACCCCTCGGACGCGGTTCTGTTGAGGGATC 420
Db 454 TCAAGGTTTACAGAGACCAAGGTGTCTACCCCTCGGACGCGGTTCTGTTGAGGGATC 513
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCACTGCGCTCTTCTGTGACACC 480
Db 514 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCACTGCGCTCTTCTGTGACACC 573
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGGACGTCGCGGACGAGGGAGGACCTG 540
Db 574 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGGACGTCGCGGACGAGGGAGGACCTG 633
QY 541 GAGCAGATTCTGAGGACGATACACACCTTCGTAAGCGCGCTTCGAGGAGTTCTGCTG 600

|||||
Db 634 GAGAGATTCTGACGCACTACACACCTTGATGAAGCGGCTTCGAGGAGTTCTGCTG 693
Qy 601 CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 660
Db 694 CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 753
Qy 661 AACCTGATCGGAGCATCATCCAGGACATTCGTGATGATGATCATCTGCAATGAGCAGCGA 720
Db 754 AACCTGATCGGAGCATCATCCAGGACATTCGTGATGATGATGATCATCTGCAATGAGCAGCGA 813
Qy 721 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 780
Db 814 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 873
Qy 781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCGAGACCCCACTGA 834
Db 874 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCGAGACCCCACTGA 927

RESULT 3
LOCUS BD157613 2160 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2160)
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 12456 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002
PF 28-JUL-2000 JP 200280990
PI TOSHIO OTA,TAKAO ISOGAI,TERUSO NISHIKAWA,KOJI HAYASHI,KOORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TERUSUI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
1.2160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN
Query Match 100.0%; Score 834; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 2e-112;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGAGGCGACCGTCCGAC 60
Db 95 ATGGCTTCGGCGGAGCGAAGACTGGAGAGCCCGCGCGGAGGCGACCGTCCGAC 154
Qy 61 CACCGGCGCTTCTGATAGGGGTGAGCGGCGCACTGCGACGGGAAAGTGCACCTGTGT 120
Db 155 CACCGGCGCTTCTGATAGGGGTGAGCGGCGCACTGCGACGGGAAAGTGCACCTGTGT 214
Qy 121 GAGAAATCATGAGTTGTGCGGACAGAAAGAGTGAACAGCGGCGCGGAAGTGTGTC 180

|||||
Db 215 GAGAAATCATGAGTTGTGCGGACAGAAAGAGTGAACAGCGGCGCGGAAGTGTGTC 274
Qy 181 ATCTTAGCCAGGACAGGTTCTTACAAAGTCTTGAACGCGCAGAGCAAGAGCGCAAGCTTG 240
Db 275 ATCTTAGCCAGGACAGGTTCTTACAAAGTCTTGAACGCGCAGAGCAAGAGCGCAAGCTTG 334
Qy 241 AAGGACAGTACAAATTTTGAACCATTCAGATGCTTTGATTAATGATTTGATGACAGAGCT 300
Db 335 AAGGACAGTACAAATTTTGAACCATTCAGATGCTTTGATTAATGATTTGATGACAGAGCT 394
Qy 301 CTGAAGAAATCGTGGAGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
Db 395 CTGAAGAAATCGTGGAGGGCAAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 454
Qy 361 TCAAGTTACACAGACCAAGGTGTTCTACCTCGGACGTTGTTTGAAGGCACTC 420
Db 455 TCAAGTTACACAGACCAAGGTGTTCTACCTCGGACGTTGTTTGAAGGCACTC 514
Qy 421 TTGGTTTCTACAGCAGAGGATCCGGGACATGTTCCACTGGCTCTTCGTGACACC 480
Db 515 TTGGTTTCTACAGCAGAGGATCCGGGACATGTTCCACTGGCTCTTCGTGACACC 574
Qy 481 GACTCCGAGCTGAGCTGTCTCGAAGAGTTCTCCGGAAGTGGCGGAGGAGGACCTG 540
Db 575 GACTCCGAGCTGAGCTGTCTCGAAGAGTTCTCCGGAAGTGGCGGAGGAGGACCTG 634
Qy 541 GAGCAGATTCTGACGCGAGTACACCACTTCTGTAAGCCGGCTTCGAGGAGTTCTGCTG 600
Db 635 GAGCAGATTCTGACGCGAGTACACCACTTCTGTAAGCCGGCTTCGAGGAGTTCTGCTG 694
Qy 601 CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 660
Db 695 CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGCACAAATATGTTGCCATC 754
Qy 661 AACCTGATCGTGCAGACATCCAGGACATTTGATGATGATGATGATGATGATGATGATGATGAT 720
Db 755 AACCTGATCGTGCAGACATCCAGGACATTTGATGATGATGATGATGATGATGATGATGATGAT 814
Qy 721 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 780
Db 815 GAGGGGTCCATGGGCGGAGCTACAGCGGACCTTTCTGAGCGGAGGACCACTCTGGG 874
Qy 781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCGAGACCCCACTGA 834
Db 875 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCGAGACCCCACTGA 928

RESULT 4
LOCUS AK022317 2160 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens CDNA FLJ12255 f16, clone MAMMA1001476, highly similar
to URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping; f16 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Matanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahashi,K., Masuo,Y., Nimomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
REFERENCE
JOURNAL NEDO human cDNA sequencing project
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission

JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - & 3' - end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
source
Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA101476"
/issue_type="Mammary gland"
/clone_lib="MAMMA1"
/note="Cloning vector: pME18SFL3"
95..928
/note="unamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASAGEDCEPAPADRPHPRLIGVSGTASGKSTCEKIM ELIGONEVORORRVVILISODRFYKVLTAEGKAKALGQVNDHPDFNDLHRLTK NIVGKTEVEPTYPYVTHSRLEPETVVPADVLPFGILVPSOEIRDMHRLPVD DSDVRLSRVLRVDRGRDLBOILQYTFVFPAPFEFCLEPKTKADVILIPGVDMV AINLIYQHIQDILNIDICMHRGSGNSRYKTFSPGDHPMLTSGKSHLESSRP H"

CDS
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN

Query Match 100.0%; Score 834; DB 9; Length 2160;
Best Local Similarity 100.0%; Pred. No. 2e-172;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGGCGAAGATCGGAGAGCCCGCGCGGAGCCGTCGCCAC 60
95 ATGGCTTCGGCGGAGGCGAAGATCGGAGAGCCCGCGCGGAGCCGTCGCCAC 154
61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGGAAAGTCGACCGTGT 120
155 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGGAAAGTCGACCGTGT 214
121 GAGAAGATCATGGAGTCTGGGACAGAACAGAGTGGAAACGCGCAGCGGAAGTGTCTC 180
215 GAGAAGATCATGGAGTCTGGGACAGAACAGAGTGGAAACGCGCAGCGGAAGTGTCTC 274
181 ATCCGAGCCGAGGACAGGTTCTACAAAGTCTGACGGGAGAGCAAGAGCCCAAGCCCTTG 240
275 ATCCGAGCCGAGGACAGGTTCTACAAAGTCTGACGGGAGAGCAAGAGCCCAAGCCCTTG 334
241 AAAGAGCAGTCAATTTTTCATCCAGATCCTTTGATTAATGATTTGATGACAGAGACT 300
335 AAAGAGCAGTCAATTTTTCATCCAGATCCTTTGATTAATGATTTGATGACAGAGACT 394
301 CTGAAGAACATCGTGGAGGCAAAAACGTTGAGAGTGCCTTATGATTTTGTACACAC 360
395 CTGAAGAACATCGTGGAGGCAAAAACGTTGAGAGTGCCTTATGATTTTGTACACAC 454
361 TCAAGGTTACGAGAGACCAAGGTTCTACCTCGGAGCGGAGTTCGTTTGAAGGGATC 420
455 TCAAGGTTACGAGAGACCAAGGTTCTACCTCGGAGCGGAGTTCGTTTGAAGGGATC 514
421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGAGACC 480
515 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGAGACC 574
481 GACTCCGACGTCAGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGGAGGAGCCTG 540
575 GACTCCGACGTCAGGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGGAGGAGCCTG 634

QY 541 GAGCAGATTCTGACGAGTACACCACTTTCGTGAAGCCGACCTTCGAGGAGTTCTGCTG 600
DB 635 GAGCAGATTCTGACGAGTACACCACTTTCGTGAAGCCGACCTTCGAGGAGTTCTGCTG 634
QY 601 CCGACAAAGAGTATGCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 660
DB 695 CCGACAAAGAGTATGCCGATGATCATCCACGAGAGTGGACAATATGTTGCCATC 754
QY 661 AACCTGATCGTGAACATCCAGGACATTCGAATGAGTGAACATTCGAAATGGACCGA 720
DB 755 AACCTGATCGTGAACATCCAGGACATTCGAATGAGTGAACATTCGAAATGGACCGA 814
QY 721 GAGAGGTTCAATGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGCAACCTGGG 780
DB 815 GAGAGGTTCAATGGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGCAACCTGGG 874
QY 781 ATGCTGACCTTGGCAACCGTCACTTTGAGTCCAGAGCAGACCCCACTGA 834
DB 875 ATGCTGACCTTGGCAACCGTCACTTTGAGTCCAGAGCAGACCCCACTGA 928

RESULT 5
AF254133
LOCUS Homo sapiens uridine kinase mRNA, complete cds. PRI 02-MAY-2001
DEFINITION AF254133
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 834)
AUTHORS Ho, Y.S. and Johnson, R.K.
TITLE Human uridine kinase from prostate cancer cell line (LNCap)
REFERENCE 2 (bases 1 to 834)
AUTHORS Ho, Y.S. and Johnson, R.K.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2000) Oncology Research, Smithkline Beecham, 709 Swedeland Road, King of Prussia, PA 19406, USA

FEATURES
source
1..834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAK49122.1"
/db_xref="GI:13924750"

CDS
BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN

Query Match 99.8%; Score 832.4; DB 9; Length 834;
Best Local Similarity 99.8%; Pred. No. 4.7e-172;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGGCGAAGATCGGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
DB 1 ATGGCTTCGGCGGAGGCGAAGATCGGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCGACCGTGTGT 120

```

QY 121 GAGAAATCATGAGTGTCTGGAGCAGAACGAGGTGGAAACGAGCGGCGAGGAGTGTCTC 180
    |||
Db 121 GAGAAATCATGAGTGTCTGGAGCAGAACGAGGTGGAAACGAGCGGCGAGGAGTGTCTC 180
QY 181 ATCTTGAGCCAGGACGAGTCTCAACAGTCTCTGACGAGCAGAGCAGAAAGGCCAAGCCTTG 240
    |||
Db 181 ATCTTGAGCCAGGACGAGTCTCAACAGTCTCTGACGAGCAGAGCAGAAAGGCCAAGCCTTG 240
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTTGATATATGATTTGATGACAGACT 300
    |||
Db 241 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTTGATATATGATTTGATGACAGACT 300
QY 301 CTGAAAGAACATCGTGAGGGGCAAAACGATGGAGGTGCCAGCTTATGATTTTGTGACAC 360
    |||
Db 301 CTGAAAGAACATCGTGAGGGGCAAAACGATGGAGGTGCCAGCTTATGATTTTGTGACAC 360
QY 361 TCAAGGTTACAGAGACCAACGATGCTCACTGCGGACGATGATTTGTTGAGGAGCATC 420
    |||
Db 361 TCAAGGTTACAGAGACCAACGATGCTCACTGCGGACGATGATTTGTTGAGGAGCATC 420
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTTGACACC 480
    |||
Db 421 TTGGTGTCTTACAGCCAGGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTTGACACC 480
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGAGGAGCCTG 540
    |||
Db 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGAGGAGCCTG 540
QY 541 GAGCAGATTCTGACGACGATACACACCTTGTAAGCCGAGCTTTCAGAGATTCTGCTG 600
    |||
Db 541 GAGCAGATTCTGACGACGATACACACCTTGTAAGCCGAGCTTTCAGAGATTCTGCTG 600
QY 601 CCGACAAAGAAATGTGCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 660
    |||
Db 601 CCGACAAAGAAATGTGCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 660
QY 661 AACCTGATCGGAGCAGCATCCAGGACATTTGTAATGTGATCTGCAATGTGACACGA 720
    |||
Db 661 AACCTGATCGGAGCAGCATCCAGGACATTTGTAATGTGATCTGCAATGTGACACGA 720
QY 721 GAGAGGTTCATGAGCGGAGCTACCAACGAGCAGCTTTTCTGAGCCAGGGGACCACTG 780
    |||
Db 721 GAGAGGTTCATGAGCGGAGCTACCAACGAGCAGCTTTTCTGAGCCAGGGGACCACTG 780
QY 781 ATGCTGACCTCTGCGAAACGCTCACATTTGAGATGTCAGACAGACCCCACTGA 834
    |||
Db 781 ATGCTGACCTCTGCGAAACGCTCACATTTGAGATGTCAGACAGACCCCACTGA 834

RESULT 6
AF237290 1022 bp mRNA linear PRI 02-AUG-2001
LOCUS Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.
DEFINITION AF237290
ACCESSION AF237290.1 GI:13506764
VERSION AF237290.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.
TITLE Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases
JOURNAL Mol. Pharmacol. 59 (5), 1181-1186 (2001)
MEDLINE 21203813
PUBMED 11306702
REFERENCE 2 (bases 1 to 1022)
AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1022)

```

```

AUTHORS Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,
Johansson,M. and Karlsson,A.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge
University Hospital, Stockholm 14186, Sweden
FEATURES
source
1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1022
/gene="UCK1"
5..838
/gene="UCK1"
/function="phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate"
/product="uridine-cytidine kinase 1"
/protein_id="AAK28324.1"
/db_xref="GI:13506765"
/translation="MASAGEDCESPAPADPHORPFLIGVSGTSAGSKSYCEKIM
ELLGQNEVEQRQRKVILISQDRFYKVLAEQAKALKQYDFDHPAQNDFLHRTLK
NIVEGKTVFVPTDVEVTHSRLEPFTVADVLEGIIVFYSOEIRDFHRLFVDT
DSVRLSRVLRDVRGRDLFOILQYTFVAPAEFELPKRYADVII PRGVDMV
AINLIYQHIQDILNGLDICKMHRGSGNRSYKRTFSEPDHPLTSGKSHLESSRP
H"
BASE COUNT 227 a 282 c 324 g 189 t
ORIGIN
Query Match 99.8%; Score 832.4; DB 9; Length 1022;
Best local Similarity 99.9%; Pred. No. 4,7e-172;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCTTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGTCCGAC 60
    |||
Db 5 ATGGCTTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGTCCGAC 64
QY 61 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGCGGCGACGACGCGGAGAAATCGATGCT 120
    |||
Db 65 CAGCGGCCCTTCTCTGATTAAGGGGTGAGCGCGGCGACGACGCGGAGAAATCGATGCT 124
QY 121 GAGAAATCATGAGTGTCTGGAGCAGAACGAGGTGGAAACGAGCGGCGAGGAGTGTCTC 180
    |||
Db 125 GAGAAATCATGAGTGTCTGGAGCAGAACGAGGTGGAAACGAGCGGCGAGGAGTGTCTC 184
QY 181 ATCTTGAGCCAGGACGAGTCTCAACAGTCTCTGACGAGCAGAGCAGAAAGGCCCTTG 240
    |||
Db 185 ATCTTGAGCCAGGACGAGTCTCAACAGTCTCTGACGAGCAGAGCAGAAAGGCCCTTG 244
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTTGATATATGATTTGATGACAGACT 300
    |||
Db 245 AAAGGACAGTACATTTTGAACATCCAGATGCCCTTTGATATATGATTTGATGACAGACT 304
QY 301 CTGAAAGAACATCGTGAGGGGCAAAACGATGGAGGTGCCAGCTTATGATTTTGTGACAC 360
    |||
Db 305 CTGAAAGAACATCGTGAGGGGCAAAACGATGGAGGTGCCAGCTTATGATTTTGTGACAC 364
QY 361 TCAAGGTTACAGAGACCAACGATGCTCACTGCGGACGATGATTTTGTGAGGAGCATC 420
    |||
Db 365 TCAAGGTTACAGAGACCAACGATGCTCACTGCGGACGATGATTTTGTGAGGAGCATC 424
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTTGACACC 480
    |||
Db 425 TTGGTGTCTTACAGCCAGGAGATCCGAGACATGTTCCACTGCGCTCTTCTGTTGACACC 484
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGAGGAGCCTG 540
    |||
Db 485 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGAGGAGCCTG 544
QY 541 GAGCAGATTCTGACGACGATACACACCTTGTGAAGCCGCGCTTTCAGAGAGTTCTGCTG 600
    |||
Db 545 GAGCAGATTCTGACGACGATACACACCTTGTGAAGCCGCGCTTTCAGAGAGTTCTGCTG 604

```

```

QY      601 CCGACAAAGAGTATGCCGATGTGATCATCCACGAGAGTGA CAATATGTTGCCATC 660
Db      605 CCGACAAAGAGTATGCCGATGTGATCATCCGCGAGAGTGA CAATATGTTGCCATC 664
QY      661 AACCTGATCTGCGACGACATCCAGGACATTCTGAATGTTGCAATCTGCACCCGA 720
Db      665 AACCTGATCTGCGACGACATCCAGGACATTCTGAATGTTGCAATCTGCACCCGA 724
QY      721 GGAGGGTCCATGAGCGGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACACCCCTGGG 780
Db      725 GGAGGGTCCATGAGCGGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACACCCCTGGG 784
QY      781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCAGACAGCCCACTGA 834
Db      785 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCAGACAGCCCACTGA 838

RESULT 7
AF125106      1395 bp  mRNA  linear  PRI 07-FEB-2002
LOCUS      Homo sapiens uridine kinase mRNA, complete cds.
DEFINITION
ACCESSION      AF125106
VERSION      AF125106.1 GI:18568108
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1395)
AUTHORS      Xun,Y.R., Yu,L. and Zhao,S.Y.
TITLE      Cloning of a new human cDNA similar to Mus musculus uridine kinase
              mRNA
JOURNAL      Unpublished
RECORD      2 (bases 1 to 1395)
AUTHORS      Ding,J.B., Yu,L. and Zhao,S.Y.
TITLE      Direct Submission
JOURNAL      Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
              Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
              People's Republic of China
FEATURES
source      1..1395
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             8..838
             /codon_start=1
             /product="uridine kinase"
             /protein_id="AA175943.1"
             /db_xref="GI:18568109"
             /translation="MASAGGEGARARAGANRPHQRPLIGVGGTASGKSTVEKIME
              LIGONEVEORHAKVILISDRFYKVLTAEGKAKLKGQYNDHDAEDMDLMTLKN
              IVEGKTEVEPTVDSRLPETVTVPADVLEGGILVDSOEIRDMPLRLFDVTD
              SDVRLSRVRLADVRGRDLQIILQYTFPKPAPEEBCPLTKYADVIRGVDNRYA
              INLIVHIQDILANDICTKMHGSGNGRTYKTFSEPDHPDMLTSGKRSHLESSRPH
              "
CDS
             /product="uridine kinase"
             /protein_id="AA175943.1"
             /db_xref="GI:18568109"
BASE COUNT      325 a      352 c      411 g      307 t
ORIGIN
Query Match      94.2%; Score 786; DB 9; Length 1395;
Best Local Similarity 97.2%; Pred. No. 6,9e-162;
Matches 811; Conservative 0; Mismatches 20; Indels 3; Gaps 1;

QY      1 ATGCTTCTGCGGAGGAGCGAAGATCTGCGAGACCCCGGCGGAGCGACCGTCCGAC 60
Db      8 ATGCTTCTGCGGAGGAGCGAAG--GTGCGAGAGCCCGCGGAGGAAACGTCGCCAC 64
QY      61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCGTGTGT 120
Db      65 CAGCGTCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCGTGTGT 124
QY      121 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGCGGACCGGAGAGTGTCT 180
Db      125 GAGAAATCATGAGTGTCTGGGACAGAAAGAGTGAACAGCGCGGACCGGAAAGTGTCT 184

```

```

QY      181 ATCTGAGCCAGGACAGAGTTCTACAAAGTCTGACGCGACAGAGGACCAAGCCCTTG 240
Db      185 ATCTGAGCCAGGACAGAGTTCTACAAAGTCTGACGCGACAGAGGACCAAGCCCTTG 244
QY      241 AAAGACAGTACATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGACT 300
Db      245 AAAGACAGTACATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGACT 304
QY      301 CTGAAGAACATCTGTGAGGGGCAAAACGTCGAGTCCGACCTATGATTTTGTGACAC 360
Db      305 CTGAAGAACATCTGTGAGGGGCAAAACGTCGAGTCCGACCTATGATTTTGTGACAC 364
QY      361 TCAAGTTACAGAGACCAACGCTGCTCTACCTCGCGACGTCGTTCTGTTTGGAGGCATC 420
Db      365 TCAAGTTACAGAGACCAACGCTGCTCTACCTCGCGACGTCGTTCTGTTTGGAGGCATC 424
QY      421 TTGGTGTCTACAGCCAGAGATCCGGAGATTTCCACTGCGCTCTTCTGTGACACC 480
Db      425 TTGGTGTCTACAGCCAGAGATCCGGAGATTTCCACTGCGCTCTTCTGTGACACC 484
QY      481 GACTCCGACGTCAGCTGTCTCGAAGATTCCTCGGAGCGTGGCCGAGGAGGACCTG 540
Db      485 GACTCCGACGTCAGCTGTCTCGAAGATTCCTCGGAGCGTGGCCGAGGAGGACCTG 544
QY      541 GAGCAGATTCTGACGACGATACCACTTCGTGAGCGGACCTTCGAGAGATTCTGCTG 600
Db      545 GAGCAGATTCTGACGACGATACCACTTCGTGAGCGGACCTTCGAGAGATTCTGCTG 604
QY      601 CCGACAAAGAGTATGCCGATGTGATCATCCACGAGAGTGA CAATATGTTGCCATC 660
Db      605 CCGACAAAGAGTATGCCGATGTGATCATCCGCGAGAGTGA CAATATGTTGCCATC 664
QY      661 AACCTGATCTGCGACGACATCCAGGACATTCTGAATGTTGCAATCTGCACCCGA 720
Db      665 AACCTGATCTGCGACGACATCCAGGACATTCTGAATGTTGCAATCTGCACCCGA 724
QY      721 GGAGGGTCCATGAGCGGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACACCCCTGGG 780
Db      725 GGAGGGTCCATGAGCGGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACACCCCTGGG 784
QY      781 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCAGACAGCCCACTGA 834
Db      785 ATGCTGACCTCTGCGAAACGGTCACTTTGGAGTCCAGCAGACAGCCCACTGA 838

RESULT 8
AK057848      2228 bp  mRNA  linear  PRI 31-OCT-2001
LOCUS      Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
              URIDINE KINASE (EC 2.7.1.48).
DEFINITION
ACCESSION      AK057848
VERSION      AK057848.1 GI:16553809
KEYWORDS      oligo capping, fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS      Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
              Fukuzumi,Y., Fujimori,Y., Komiyaama,M., Suzuki,Y., Hata,H.,
              Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T.,
              Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
              Kawakami,B., Nagai,K., Isogai,T. and Sugano,S.
              NEDO human cDNA sequencing project
              Unpublished
              2 (bases 1 to 2228)
              Sugano,S. and Suzuki,Y.
              Direct Submission
              Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
              University of Tokyo, Laboratory of Genome Structure, Human Genome
              Center, Shitokane-dai, 4-6-1, Minato-Ku, Tokyo 108-8639, Japan
              (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,

```


Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 61 Row: P Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
 source
 1..1959
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:36231 IMAGE:4913412"
 /issue_type="Salivary gland, 10 week old female mouse"
 /clone_lib="NCI CGAP_Sg2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

gene

CDS

1..1959
 /gene="Umpk"
 /db_xref="locusid:22245"
 /db_xref="MGI:98904"
 57..890
 /codon_start=1
 /product="Umpk protein"
 /protein_id="AAH25146.1"
 /db_xref="GI:19263564"
 /db_xref="locusid:22245"
 /translation="MASAGGSGSSAARPPPLIGVSGTASGKSTVCEKTM ELIGQNEVDRQRKVLISQCFKVLAEQKALAGQVNFDPDAFNDLHKITLK NIVGKTEVETPDVFTVHSRLPETVVPAPVAFBEILVEYDEIDMHLRLFDVT DSDVRLSRVLRDVRGRDLLEILTOYTAFFKPAFEFCPTKXAVIIPRGVDNV AINLVGHIDILNGDLCRRHRGPPNRNKHRTPEPDHGVATLAKSHLSSSRP H"

BASE COUNT 466 a 511 c 561 g 421 t
 ORIGIN

Query Match 78.9%; Score 658; DB 10; Length 1959;
 Best Local Similarity 86.8%; Pred. No. 7.7e-134;
 Matches 724; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGGAAGTCTGCGAGAGCCCGCGCGGAGCGGACCTCCGCAC 60
 Db 57 ATGGCTTCGGCGGAGGCGGCGGCTCGAGAGCCCGCGCGGAGCGGACCTCCGCAC 116
 QY 61 CAGGGGCGCTTCGATAGGGGTGAGGGGCGGCACTCGACGGGAGTCCAGCTGTGT 120
 Db 117 CCGGGCGGCTTCATCGCGGTAGCGGCGGCACTGATGCAAGTCAACGTGTGT 176
 QY 121 GAGAGATCATGAGTCTGCGAGAGAGAGGTGGAACAGCGGAGCGGAGAGTGTGTC 180
 Db 177 GAGAGATCATGAGTCTGCGAGAGAGAGTGAACAGCGGAGCGGAGAGTGTGTC 236
 QY 181 ATCTGAGCCAGGAGAGGTTCTACAGGTTCTGACGCGAGAGAGAGGCGGCTTG 240
 Db 237 ATCTGAGCCAGGAGAGGTTCTACAGGTTCTGACGCGGAGAGAGGCGGAGGCTTTG 296
 QY 241 AAAGAGATGATATTTTGAACATCCAGATGCTTATTAATGATTTGATGCAAGAGACT 300
 Db 297 AAAGAGATGATATTTTGAACATCCAGATGCTTATTAATGATTTGATGCAAGAGACT 356
 QY 301 CTGAGAGATCATGAGAGGCGGAGAGAGGAGTGCAGACTATGATTTTGTGACACAC 360
 Db 357 CTGAGAGATCATGAGAGGCGGAGAGAGGAGTGCAGACTATGATTTTGTGACACAC 416
 QY 361 TCAAGGTATCCAGAGACCGGTGTCTACCTTCGCGAGCTGTGTTGTTGAGGGATC 420
 Db 417 TCAAGGTATCCAGAGACCGGTGTCTACCTTCGCGAGCTGTGTTGTTGAGGGATC 476
 QY 421 TTGGTGTCTACAGGAGAGAGATCCGGAGATGTTTCAACCTGCGGCTCTTCTGTGGAACC 480
 Db 477 TTGGTGTCTACAGGAGAGAGATCCGGAGATGTTTCAACCTGCGGCTCTTCTGTGGAACA 536

QY 481 GACTCCAGCTCAGGCTGTCTCGAAGATTCTCCGGAGACGTGCGCGGAGGAGGACCTG 540
 Db 537 GACTCTGATTTAGGCTGTCTCGAAGATTCTCCGGAGATGCAAGAGAGAGGACCTG 596
 QY 541 GAGCAGATTCTGACGAGTACACACCTTCTGTGAAGCCGGCTTTCGAGAGTTCGTGCTG 600
 Db 597 GACGAGATCTGACGAGTACACACCTTCTGTGAAGCCGGCTTTCGAGAGTTCGTGCTG 656
 QY 601 CCGAGCAAGAGATGCGGATGATCATCCGAGAGAGTGTGACATATGTTGCTATC 660
 Db 657 CCGAGCAAGAGATGCGGATGATCATCCGAGAGAGTGTGATATATATGTTGCTATC 716
 QY 661 AACCTGCTGCGAGACATCCAGACATTTGAAATGTGACATCTGCAATGSCACCGA 720
 Db 717 AACCTGCTGCGAGACATCCAGACATTTGAAATGTGACATCTGCAATGSCACCGA 776
 QY 721 GAGAGGTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCGGAGGACCACTGGG 780
 Db 777 GCGGGGCCCAACGGGGCAACCAAGAGGACCTTCCCGAGCGAGATCACCTCGGG 836
 QY 781 ATGCTGACCTTGGCAAGCGTCACTTGGAGTCCAGAGGAGCCCACTGA 834
 Db 837 GTGTTGGCACTGGCAAGCGTCACTGAGTCTTACAGAGACCCCACTGA 890

RESULT 10
 BC015547
 LOCUS
 DEFINITION
 Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9658
 IMAGE:3845821, mRNA, complete cds.
 BC015547
 VERSION
 BC015547.1 GI:15930229
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2072)
 Strausberg, R.
 Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabos-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: F Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES

source

CDS

1..2072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9658 IMAGE:3845821"
 /issue_type="Colon, adenocarcinoma"
 /clone_lib="NIH_MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 87..692


```

/codon_start=1
/product="similar to uridine-cytidine kinase 1"
/protein_id="AAH15547.1"
/db_xref="GI:15930230"
/translation="MASAGGDCESPAPDPRPQRPFLIGVSGTASGSKSTCEKIM
ELLGONEVEORQRKVLISDRPFKVLTAEGKAALGGVNFDPDAPNDIMLRITLK
NIVGKNEVEPTVTPVTHSRLEPTTVAPDVVLTFCGLIYPSGEISIDMTHLRFLVDT
DSVDRLSRDREVCRCDHPTSSGGYGHQDPRAAHPRGSHK"
BASE COUNT
ORIGIN
455 a      559 c      637 g      421 t

```

Query Match	76.0%	Score 634	DB 9	Length 2072
Best Local Similarity	88.6%	Pred. No.	1.4e-128	
Matches 739	Conservative 0	Indels	95	Gaps 1

QY	1	ATGCTTTCGCGCGGAGCGAAAGACTGCGAGAGCCCGCGCCGAGGCGCAACGCTCGCAC	60
Db	87	ATGCTTTCGCGCGGAGCGAAAGACTGCGAGAGCCCGCGCCGAGGCGCAACGCTCGCAC	146
QY	61	CAGCGGCCCTTCCTGATAGGGGTGACGCGCGGCACTGCAACGGGAAGTGCACCTGTGT	120
Db	147	CAGCGGCCCTTCCTGATAGGGGTGACGCGCGGCACTGCAACGGGAAGTGCACCTGTGT	206
QY	121	GAGAAAGTCAATGAGTTCGGGACAGAACAGAGTGGAAACAGCGGACGGAAAGTGTCTC	180
Db	207	GAGAAAGTCAATGAGTTCGGGACAGAACAGAGTGGAAACAGCGGACGGAAAGTGTCTC	266
QY	181	ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGACAGCAGAAAGCCAGGCTTG	240
Db	267	ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGACAGCAGAAAGCCAGGCTTG	326
QY	241	AAAGACAGTACATTTTGAACCAACCGAATGCCCTTTGATATGATTTGATGACAGACT	300
Db	327	AAAGACAGTACATTTTGAACCAACCGAATGCCCTTTGATATGATTTGATGACAGACT	386
QY	301	CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGTGCGACCTTATGATTTTGTGACAC	360
Db	387	CTGAAGAACATCTGAGAGGCGAAACCGGTGAGGTGCGACCTTATGATTTTGTGACAC	446
QY	361	TCAAGGTTTCCAGAGACCAACGCTGCTTACCTTCGACGACGCTGTCTGTTTGAAGGATC	420
Db	447	TCAAGGTTTCCAGAGACCAACGCTGCTTACCTTCGACGACGCTGTCTGTTTGAAGGATC	506
QY	421	TTGGTGTTCACACCCAGAGAGATCCGCGGACATGTTTCCACCTGCGCCCTTCGTGGAAC	480
Db	507	TTGGTGTTCACACCCAGAGAGATCCGCGGACATGTTTCCACCTGCGCCCTTCGTGGAAC	566
QY	481	GACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGAGACGTCGCGCAGAGGAGGACCTG	540
Db	567	GACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGAGACGTCGCGCAGAGGAGGACCTG	593
QY	541	GAGCAGATTCTGACGCGACGTACACCACTTCGTGAAGCCGCGCTTCGAGAGTCTGCTTG	600
Db	594	-----	593
QY	601	CCGACAAAGAGTATGCGCATGTGATATCCCAACGAGAGTGGACAAATATGTTGGCATC	660
Db	594	--GACAAAGAGTATGCGCATGTGATATCCCAACGAGAGTGGACAAATATGTTGGCATC	651
QY	661	AACCTGATCGTGACGACATTCAGAGCATTTCTGATGTTGACATCTGCAATATGGACCGA	720
Db	652	AACCTGATCGTGACGACATTCAGAGCATTTCTGATGTTGACATCTGCAATATGGACCGA	711
QY	721	GGAGGGTTCATAGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGTGG	780
Db	712	GGAGGGTTCATAGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACCTGTGG	771
QY	781	ATGCTGACCTCTGGCAAAACGGTCACTTTTGGAGTCCAGACGACGACCCCACTGA	834
Db	772	ATGCTGACCTCTGGCAAAACGGTCACTTTTGGAGTCCAGACGACGACCCCACTGA	825

BD146824	BD146824	753 bp	DNA	linear	PAT 17-JAN-2003
LOCUS	BD146824	753 bp	DNA	linear	PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD146824				
VERSION	BD146824.1	GI:27852582			
KEYWORDS	JP 2002191363-A/1667.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Ota,T., Iwagata,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A 1667 09-JUL-2002;				
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/1667 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..753 FT /organism='Homo sapiens (human)'. Location/Qualifiers 1..753 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" BASE COUNT 158 a 196 c 261 g 135 t 3 others ORIGIN				
Query Match	74.8%	Score 623.6;	DB 6;	Length 753;	
Best Local Similarity	97.7%;	Pred. No. 2.8e-126;			
Matches	642;	Conservative	0;	Mismatches 12;	Indels 3;
					Gaps 1;
QY	1	ATGGCTTGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCCGAC	60		
DB	95	ATGGCTTGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCCGAC	154		
QY	61	CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCGACTGCGCGGAGGCGAGCGTGTGT	120		
DB	155	CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCGACTGCGCGGAGGCGAGCGTGTGT	214		
QY	121	GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTC	180		
DB	215	GAGAAATCATGAGATTGCTGGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGTC	274		
QY	181	ATCTGAGCCGAGCAGGTTTAAAGGTCTTACCGCGACGACGAGAGGCCAAGGCTTG	240		
DB	275	ATCTGAGCCGAGCAGGTTTAAAGGTCTTACCGCGACGAGAGGCCAAGGCTTG	334		
QY	241	AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGAGCT	300		
DB	335	AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGAGCT	394		
QY	301	CTGAAGAACATCGTGAAGGCGAAACGATGAGGTGCGACTTAATTTTGTGACACAC	360		
DB	395	CTGAAGAACATCGTGAAGGCGAAACGATGAGGTGCGACTTAATTTTGTGACACAC	454		
QY	361	TCAAGGTTTACGAGACCAAGGTGTCTACCTTGCGGACGTGTGTTCTGTTAGGGCATC	420		
DB	455	TCAAGGTTTACGAGACCAAGGTGTCTACCTTGCGGACGTGTGTTCTGTTAGGGCATC	514		
QY	421	TTGGTGTTCACGACCGAGGATCGGGGACATGTTTCACTGCGGCTCTTCTGTGACAC	480		

```

Db      515 TTGGTGTTCACACCGAGGATCCGGGACATGTTCCACCTCCGCTCTTGTGTGACACC 574
Qy      481 GACTCCGACGTCAGCGTCTTCGAAAGTTCTCCGGGACGTGCCCGAAGGAGGACCTG 540
Db      575 GACTCCGACGTCAGCGTCTTCGAAAGTTCTCCGGGACGTGCCCGAAGGAGGACCTG 634
Qy      541 GAGGAGATTCGAGCGAGTACACACCTTCGTAAGCGCGCTTCGAGGAGTTCTGCC-- 598
Db      635 GAGGAGATTCGAGCGAGTACACACCTTCGTAAGCGCGCTTCGAGGAGTTCTGCCCT 694
Qy      599 -TGCCGCAAAAGAGTATGCCGATGTGATCATCCACAGAGGTGACAATATGTT 654
Db      695 TGCGCAGCAAAAGAGTATGCCGATGTGATCATTCACAGAGGTGACAAATATGNT 751

RESULT 12
MUSURKI      1810 bp  mRNA  linear  ROD 27-MAR-1997
LOCUS      Mus musculus uridine kinase mRNA, partial cds.
DEFINITION  L31783
ACCESSION  L31783.1 GI:471980
VERSION    L31783.1
KEYWORDS
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1810)
AUTHORS    Traut,T.W.
TITLE      The functions and consensus motifs of nine types of peptide
           segments that form different types of nucleotide-binding sites
JOURNAL    Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE    8200357
PUBMED     2 (bases 1 to 1810)
REFERENCE   Ropp,P.A. and Traut,T.W.
AUTHORS    Cloning and expression of a cDNA encoding uridine kinase from mouse
           brain
JOURNAL    Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
MEDLINE    97108719
PUBMED     8951040
FEATURES
source     location/Qualifiers
           1..1810
           /organism="Mus musculus"
           /mol_type="mRNA"
           /db_xref="taxon:10090"
           /tissue_type="Brain"
           <1..783
           /EC_number="2.7.1.48"
           /standard_name="uridine-cytidine kinase"
           /codon_start=1
           /product="uridine kinase"
           /protein_id="AAB50568.1"
           /db_xref="GI:471981"
           /translation="RPPRPFLIGSGTASGKSTVEKIMELLIGONEVDRRORLVI
           LSOCEYKVLTAEOKAKLGOVNFDPDAPNDIMKTKLKNIVEGTVAVPYDPVT
           HSRLEPTVVYPADIVLPEGLIVYKQIODEIMELRLFVPTDSSVRLSRVLDVQNG
           RDLFOILTYTAFYKPAFEBCFLPTKXADVILPRGVNVALILVQHTQIDILNGDL
           CKHRGGPNGRNHRKTPPEPGDHGVATGKRSLSRPH"
           join(37..63,412..426,451..465)
           /bound_molety="ATP at catalytic site"
           misc_binding
           polyA_signal 1790..1795 512 g 405 t
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN
Query Match 74.3%; Score 619.8; DB 10; Length 1810;
Best Local Similarity 87.0%; Pred.No.1.8e-125;
Matches 681; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy      52 CGTCCGACACGCGGCTTCTCTATAGGGGTGAGCGCGGACATGCCAGCGGGAAGTCG 111
Db      1 CGTCCGACGCGCGGCTTCTCTATAGGGGTGAGCGCGGACATGCCAGCGGGAAGTCG 60

```

```

Qy      112 ACCGTGTGTAAGAGATCATGAGTTGCTGGGACAGAACGAGTGGAGACGCGACGG 171
Db      61 ACAGTGTGTAAGAGATCATGAGTTGCTGGGACAGAACGAGTGGAGACGCGACGG 120
Qy      172 AAGTGTGATCTCTGAGCCAGAGCAGGTTCTACAGAGTCTTGAACGAGCAAGAGCC 231
Db      121 AAGTGTGATCTCTGAGCCAGAGCAGTCTTCTACAGAGTCTTGAACGAGCAAGAGCC 180
Qy      232 AAGGCTTGAAGAGACATGATTTTGAACATCCAGATGCCATGATATATATTTGATG 291
Db      181 AAGGCTTGAAGAGACATGATTTTGAACATCCAGATGCCATGATATATATGATGATG 240
Qy      292 CACAGGACTCTGAAGACATCGTGAAGGCAAAACGGTGAAGGTGCGACCTATGATTTT 351
Db      241 CACAGGACTCTGAAGACATCGTGAAGGCAAAACGTGAGGTCCCTTACTATGATTTT 300
Qy      352 GTGACACTCAAGTTACACAGACACGAGTGTCTACCTCGGACGAGTGTCTGTTT 411
Db      301 GTGACCACTCAAGTTACACAGACCACTGTGTCTACCCAGCTGATGTGTTCTGTTT 360
Qy      412 GAGGCACTGTGTTCTTACAGCAGAGATCCGAGATGTCACCTGCTGCTTC 471
Db      361 GAGGCACTGTGTTCTTACAGCAGAGATCCGAGATGTCACCTGCTGCTTC 420
Qy      472 GTGACACTCACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGAGG 531
Db      421 GTGACACTCACTCTGATGTAGGCTGTCTCGAAGAGTTCTCCGGATGTGACAGAGA 480
Qy      532 AGGAGCTTGAGAGAGATTTGACGAGTACACCACTTCGTGAAGCGGCTTTCAGAGG 591
Db      481 AGGAGCTTGAGAGAGATTTGACGAGTACACCACTTCGTGAAGCGGCTTTCAGAGG 540
Qy      592 TTCTGCTGCGGCAAAAGAGTATGCCGATGTATATATCCACAGAGATGTGACATATG 651
Db      541 TTCTGCTGCGGCAAAAGAGTATGCCGATGTATATATCCACAGAGATGTGACATATG 600
Qy      652 GTTGCCATCAACTGATGTCGACGACATTCAGAGACATTCGTGAATGTGACATTCGAAA 711
Db      601 GTTGCCATCAACTGATGTCGACGACATTCAGAGACATTCGTGAATGTGACATTCGAAA 660
Qy      712 TGCGACCGAGAGAGGTTCATGCGGAGCTACAGACGACCTTTTCTGAGCCAGGAG 771
Db      661 CGGACCGAGAGAGGTTCATGCGGAGCTACAGACGACCTTTTCTGAGCCAGGAGAT 720
Qy      772 CACCTTGAGATGTCGACCTCTGCAACAGGTGACATTTGAGATTCAGACAGACCCAC 831
Db      721 CACCTTGAGATGTCGACCTCTGCAACAGGTGACATTTGAGATTCAGACAGACCCAC 780
Qy      832 TGA 834
Db      781 TGA 783

RESULT 13
MUSURKIA      952 bp  DNA  linear  ROD 05-MAR-2001
LOCUS      Mus musculus uridine kinase pseudogene.
DEFINITION  L31784
ACCESSION  L31784.1 GI:471982
VERSION    L31784.1
KEYWORDS   pseudogene; uridine kinase.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 952)
AUTHORS    Traut,T.W.
TITLE      The functions and consensus motifs of nine types of peptide
           segments that form different types of nucleotide-binding sites
JOURNAL    Eur. J. Biochem. 222 (1), 9-19 (1994)
MEDLINE    8200357
PUBMED     2 (bases 1 to 952)
REFERENCE   Ropp,P.A. and Traut,T.W.
AUTHORS

```


TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse brain
 JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
 MEDLINE 97108719
 PUBMED 8951040
 FEATURES
 source Location/Qualifiers
 1..952
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /tissue_type="kidney"
 76..913
 /standard_name="uridine-cytidine kinase"
 /pseudo
 /codon_start=1
 /product="uridine kinase"
 misc_difference 130
 /note="insertion not found in L31783"
 BASE COUNT 230 a 231 c 277 g 214 t
 ORIGIN
 Query Match 70.9%; Score 591.2; DB 10; Length 952;
 Best Local Similarity 82.5%; Pred. No. 3.5e-119;
 Matches 691; Conservative 0; Mismatches 143; Indels 4; Gaps 1;

1 ATGGCTTGGGCGGAGGCGAAGACTGGAGGCCCCGCGC-----GCCGGAGGCCGACCGCTCC 56
 76 ATGGCTTGGGCGGAGGCGGCTGCTCAGAGAGTACGCGCTGCCGCGCGCATCTTCC 135
 57 GCAACGAGCGCCCTCTCTGATAGAGGTGAGCGGCGGCACTGCCAGCGGAGTCAACCT 116
 136 TCAGCCCTTGGCCGCTCTCATCGGCGGTGAGGGGCGGACCGCTAGTGGCAAGTCAACGT 195
 117 GTGTGAGAAATCATGAGATTGCTGGGACAGAACGAGCTGGAACAGCGGCAAGGAA 176
 196 GTGTGAGAAATCATGAGCTGCTGGGACAGAAATGAACTGACCTCCGCTAGCGCAAGTT 255
 177 GGTGATCTGAGCGAGGACAGGTTCTACAGGTCTGACGCGGACAGAGCAAGGCGAAAGC 236
 256 GGTGATCTGAGCGAGGACAGGCTCTACAGGTCTGACGCGGACAGAGCAAGGCGAAAGC 315
 237 CTTGAAAGGACAGTACATTTTTCAGATCCAGATGCTTTGATATGATTTGATGACAG 296
 316 TTGGAAGGACAGTACATTTTTCAGATGCTTTGATATGATTTGATGACAG 375
 297 GACTCTGAGAAACATCTGTGAGAGGCGAAACGCTGAGAGTCCGACCTATGATTTTGTAC 356
 376 GACCTGAGAAACATTTATGAGGCGAAAGTTGTTAAGTTCTTACCTATGATTTTGTAC 435
 357 ACACTCAAGGTTACCGAGACCAACGCTGCTACCCCTGCGGACGCTGCTTGTGAGGG 416
 436 CGACTCAAGGTTACCGAGTACCACTGCTGCTACCAATTTGATGAGCGGTTCTGTTCAGGG 495
 417 CATCTTGTTGTTTACAGCGAGAGATCCGGGACATGTTCCACCTGCGCTTCTGTCGA 476
 496 CATCTTGTTGTTTACAGCGAGATCCGGGACATGTTCCACCTGCGCTTCTGTCGA 555
 477 CACCGACTCCGAGCTCAGGCTGTCTGGAAGAGTTCTCCGGACGTCGCGCGGAGGAGGA 536
 556 CACAGACCCCTGATGTTAGGCTGTCTGGAAGAGTTCTCCGGAGTGTGCAAGAGGAGGA 615
 537 CCGGAGCAGATTTCTGACGAGTACACACCTTCGGAAGCGCGCTTCGAGAGATTCG 596
 616 CCGGAGCAGATTTCTGACGAGTACACACCTTCGGAAGCGCGCTTCGAGAGATTCG 675
 597 CCGGCGGACAAAGAGATCCGATGATCATCCGAGAGAGTGAACATATGTTGC 656
 676 CCGGCGGACAAAGAGATCCGATGATCATCCGAGAGAGTGAACATATGTTGC 735
 657 CATCAACTGATGTTGAGACATCCAGGACATTTGATGATGATTTGCAATTTGGCA 716
 736 CATCAACTGATGTTGATTAACATCCAGGACATTTCAACGAGGAGCCTGTGCAAGCGGCA 795

Qy 717 CCGAGAGGGTCCATATGAGCGGAGTACAGCGGACCTTTTCTGAGCCAGGGACCC 776
 Db 796 CCGAGCGGGGCCCAATGGGCGACGCCACAGAGACTTACCGGATCGAGATCACCC 855
 Qy 777 TGGGATGCTGACCTCTGGCAACGCGTCAATTTGAGTCCAGACGAGACCCCATGA 834
 Db 856 TGGGATGTTGGCCACTGTGTAACGCTGACACCTGAGTCTGACGACGAGACCCCATGA 913

RESULT 14
 AC078885/c
 LOCUS
 DEFINITION
 Mus musculus chromosome 6 clone RP23-117123 strain C57BL/6/J,
 WORKING DRAFT SEQUENCE, 44 unordered pieces.
 AC078885
 AC078885.8 GI:18376842
 HTG; HTGS PHASE1; HTGS DRAFT.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 254993)
 Montgomerly,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Pereira,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 High Throughput Mouse Sequencing
 Unpublished
 2 (bases 1 to 254993)
 Montgomerly,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Pereira,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.
 Direct Submission
 Submitted (08-AUG-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Jan 26, 2002 this sequence version replaced gi:14488282.
 -----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: http://www.hpcgg.org/Sequence/mouse.html
 Contact: hpcg@mednet.mgh.harvard.edu
 -----Summary Statistics
 Center project name: AAP
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye; 100%
 *Consensus quality: 247096 at least Q20
 *Consensus quality: 242518 at least Q30
 *Consensus quality: 235589 at least Q40
 Estimated insert size: agarose-PP - N/A
 **Estimated insert size: 254133 - sum-of-contigs
 Quality coverage: agarose-PP - N/A
 Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 44 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 27415: contig of 27415 bp in length
 * 27416 27435: gap of unknown length
 * 27436 57817: contig of 30382 bp in length
 * 57818 57837: gap of unknown length
 * 57838 81637: contig of 23800 bp in length
 * 81638 81657: gap of unknown length
 * 81658 102155: contig of 20498 bp in length
 * 102156 102175: gap of unknown length
 * 102176 121200: contig of 19025 bp in length
 * 121201 121220: gap of unknown length
 * 121221 135073: contig of 13853 bp in length
 * 135074 135093: gap of unknown length

135094	156689:	contig of 21596 bp in length
156590	156709:	gap of unknown length
156710	16185:	contig of 9476 bp in length
166186	166205:	gap of unknown length
166206	178798:	contig of 12593 bp in length
178799	178818:	gap of unknown length
178819	187247:	contig of 8428 bp in length
187248	187267:	gap of unknown length
187268	192492:	contig of 5225 bp in length
192493	192512:	gap of unknown length
192513	200084:	contig of 7572 bp in length
200085	200104:	gap of unknown length
200105	207054:	contig of 6950 bp in length
207055	207074:	gap of unknown length
207075	212944:	contig of 5870 bp in length
212945	212964:	gap of unknown length
212965	219468:	contig of 6504 bp in length
219469	219488:	gap of unknown length
219489	220369:	contig of 881 bp in length
220370	220389:	gap of unknown length
220390	221164:	contig of 775 bp in length
221165	221184:	gap of unknown length
221185	222796:	contig of 1612 bp in length
222797	222816:	gap of unknown length
228217	224762:	contig of 1946 bp in length
224763	224782:	gap of unknown length
224783	225999:	contig of 1217 bp in length
226000	226019:	gap of unknown length
226020	227409:	contig of 1390 bp in length
227410	227429:	gap of unknown length
227430	230145:	contig of 2716 bp in length
230146	230165:	gap of unknown length
230166	231842:	contig of 1677 bp in length
231843	231862:	gap of unknown length
231863	232786:	contig of 924 bp in length
232787	232806:	gap of unknown length
232807	233137:	contig of 331 bp in length
233138	233157:	gap of unknown length
233158	234519:	contig of 1362 bp in length
234520	234539:	gap of unknown length
234540	236416:	contig of 1877 bp in length
236417	236436:	gap of unknown length
236437	237967:	contig of 1531 bp in length
237968	237987:	gap of unknown length
237988	239220:	contig of 1233 bp in length
239221	239240:	gap of unknown length
239241	239367:	contig of 127 bp in length
239368	239387:	gap of unknown length
239388	239599:	contig of 572 bp in length
239590	239979:	gap of unknown length
239980	241016:	contig of 1037 bp in length
241017	241036:	gap of unknown length
241037	242392:	contig of 1355 bp in length
242392	242411:	gap of unknown length
242412	243823:	contig of 1412 bp in length
243824	243843:	gap of unknown length
243844	245227:	contig of 1384 bp in length
245228	245247:	gap of unknown length
245248	246334:	contig of 1087 bp in length
246335	246354:	gap of unknown length
246355	247451:	contig of 1097 bp in length
247452	247471:	gap of unknown length
247472	248780:	contig of 1309 bp in length
248781	248800:	gap of unknown length
248801	250041:	contig of 1241 bp in length
250042	250061:	gap of unknown length
250062	251173:	contig of 1112 bp in length
251174	251193:	gap of unknown length
251194	252713:	contig of 1520 bp in length
252714	252733:	gap of unknown length
252734	254014:	contig of 1281 bp in length
254015	254034:	gap of unknown length
254035	254176:	contig of 142 bp in length
135094	156689:	contig of 21596 bp in length
156590	156709:	gap of unknown length
156710	16185:	contig of 9476 bp in length
166186	166205:	gap of unknown length
166206	178798:	contig of 12593 bp in length
178799	178818:	gap of unknown length
178819	187247:	contig of 8428 bp in length
187248	187267:	gap of unknown length
187268	192492:	contig of 5225 bp in length
192493	192512:	gap of unknown length
192513	200084:	contig of 7572 bp in length
200085	200104:	gap of unknown length
200105	207054:	contig of 6950 bp in length
207055	207074:	gap of unknown length
207075	212944:	contig of 5870 bp in length
212945	212964:	gap of unknown length
212965	219468:	contig of 6504 bp in length
219469	219488:	gap of unknown length
219489	220369:	contig of 881 bp in length
220370	220389:	gap of unknown length
220390	221164:	contig of 775 bp in length
221165	221184:	gap of unknown length
221185	222796:	contig of 1612 bp in length
222797	222816:	gap of unknown length
228217	224762:	contig of 1946 bp in length
224763	224782:	gap of unknown length
224783	225999:	contig of 1217 bp in length
226000	226019:	gap of unknown length
226020	227409:	contig of 1390 bp in length
227410	227429:	gap of unknown length
227430	230145:	contig of 2716 bp in length
230146	230165:	gap of unknown length
230166	231842:	contig of 1677 bp in length
231843	231862:	gap of unknown length
231863	232786:	contig of 924 bp in length
232787	232806:	gap of unknown length
232807	233137:	contig of 331 bp in length
233138	233157:	gap of unknown length
233158	234519:	contig of 1362 bp in length
234520	234539:	gap of unknown length
234540	236416:	contig of 1877 bp in length
236417	236436:	gap of unknown length
236437	237967:	contig of 1531 bp in length
237968	237987:	gap of unknown length
237988	239220:	contig of 1233 bp in length
239221	239240:	gap of unknown length
239241	239367:	contig of 127 bp in length
239368	239387:	gap of unknown length
239388	239599:	contig of 572 bp in length
239590	239979:	gap of unknown length
239980	241016:	contig of 1037 bp in length
241017	241036:	gap of unknown length
241037	242392:	contig of 1355 bp in length
242392	242411:	gap of unknown length
242412	243823:	contig of 1412 bp in length
243824	243843:	gap of unknown length
243844	245227:	contig of 1384 bp in length
245228	245247:	gap of unknown length
245248	246334:	contig of 1087 bp in length
246335	246354:	gap of unknown length
246355	247451:	contig of 1097 bp in length
247452	247471:	gap of unknown length
247472	248780:	contig of 1309 bp in length
248781	248800:	gap of unknown length
248801	250041:	contig of 1241 bp in length
250042	250061:	gap of unknown length
250062	251173:	contig of 1112 bp in length
251174	251193:	gap of unknown length
251194	252713:	contig of 1520 bp in length
252714	252733:	gap of unknown length
252734	254014:	contig of 1281 bp in length
254015	254034:	gap of unknown length
254035	254176:	contig of 142 bp in length
FEATURES	*	254177 254196: gap of unknown length
source	254197	254993: contig of 797 bp in length.
		Location/Qualifiers
	1. .254993	
		/organism="Mus musculus"
		/mol_type="genomic DNA"
		/strain="C57BL6/J"
		/db_xref="taxon:10090"
		/chromosome="6"
		/clone="RP23-117123"
		/sex="male"
		1. .27415
		/note="assembly_name:Contig138"
		27436. .57817
		/note="assembly_name:Contig137
		clone_end:Sp6
		vector_side:left"
		57838. .81637
		/note="assembly_name:Contig136"
		81658. .102155
		/note="assembly_name:Contig135"
		102176. .121200
		/note="assembly_name:Contig134"
		121221. .135073
		/note="assembly_name:Contig133"
		135094. .156689
		/note="assembly_name:Contig132"
		156710. .166185
		/note="assembly_name:Contig131
		clone_end:T7
		vector_side:left"
		166206. .187798
		/note="assembly_name:Contig130"
		178819. .187247
		/note="assembly_name:Contig129"
		187268. .192492
		/note="assembly_name:Contig128"
		192513. .200084
		/note="assembly_name:Contig127"
		200105. .207054
		/note="assembly_name:Contig126"
		207075. .212944
		/note="assembly_name:Contig125"
		212965. .219468
		/note="assembly_name:Contig124"
		219489. .220369
		/note="assembly_name:Contig123"
		220390. .221164
		/note="assembly_name:Contig122"
		221185. .222796
		/note="assembly_name:Contig121"
		222817. .224762
		/note="assembly_name:Contig120"
		224783. .225999
		/note="assembly_name:Contig119"
		226020. .227409
		/note="assembly_name:Contig118"
		227430. .230145
		/note="assembly_name:Contig117"
		230166. .231842
		/note="assembly_name:Contig116"
		231863. .232786
		/note="assembly_name:Contig115"
		232807. .233137
		/note="assembly_name:Contig114"
Query Match	70.8%:	Score 590.4; DB 2; Length 254993;
Best Local Similarity	83.8%:	Pred. No. 3,9e-119;
Matches 654;	Conservative 0;	Mismatches 126; Indels 0; Gaps 0;
QY	55	CCGACACAGCGGCGCTTCTCGATAGGGGTGAGCGGCGGACCTGCACGCGGAATCGACC 114
DB	156715	CGGATCGNN

QY 115 GTGTGAGAAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACAGCGGCGACGGGAG 174
 Db 15655 GTGTGTAGAAAGATCATGAGCTCTGTGAGACGAAAGAGGAGCCCGCGACGGCAG 156596
 QY 175 GTGTGATCTCTGACCGACGAGCATGTTCTACAAAGTCTCTGACGCGACGACGAGAGCCGAC 234
 Db 156595 TTGGTCAATCTGACCGACGAGATGCTTCTACAAAGTCTGACGCGCGACGAGAGCCGAC 156536
 QY 235 GCCTTGAAGAGACGATCAATTTTGCACATTCAGATGCTTGTATATGATTTGATGAC 294
 Db 156535 GCCTTGAAGAGACGATCAATTTTGCACATTCAGATGCTTGTATATGATTTGATGAC 156476
 QY 295 AGGACTGTGAAGACATGTTGAGAGGAGAAACGGGTGAGAGGCGACCTATGATTTGTG 354
 Db 156475 AAGACCTGTGAAGACATGTTGAGAGGAGAAACGTGAGGTCCTTACCTATGATTTGTG 156416
 QY 355 ACACTCAAGGTATACGAGAGACCAAGGATGTTCTACCTGCGACGTGTTCTTTGAG 414
 Db 156415 ACCACTCAAGGTATACGAGAGACCAAGGATGTTCTACCTGCGATGTTCTTTGAG 156356
 QY 415 GGCACTTTGGTGTCTACAGCGAGAGATCCGGACATGTTCCACTGCGCTCTTGTG 474
 Db 156355 GGCACTTTGGTGTCTACAGCGAGAGATCCGGACATGTTCCACTGCGCTCTTGTG 156296
 QY 475 GACACGAGCTCCGACGTCAGGCTGTCTGAAAGATTTCCGGGACGTCGCGCGAGGAG 534
 Db 156295 GACACGAGCTCTGATGTTAGGCTGTCTGAAAGATTTCCGGGATGTCGACGAGAGG 156236
 QY 535 GACCTGAGAGATTTCTGACGAGATGACCACTTCTGTAAGCGCGCTTCAAGAGATTG 594
 Db 156235 GACCTGAGAGATTTCTGACGAGATGACCACTTCTGTAAGCGCGCTTCAAGAGATTG 156176
 QY 595 TGCTGCGGACAAAGATATGCCATGTGATCATCCACGAGAGTGAACATATGTT 654
 Db 156175 TGCTGCGGACAAAGATATGCCATGTGATCATCCCTGAGAGATGTTAATATGTTG 156116
 QY 655 GGCATCAACCTGATCGAGAGCATGACGAGCACTTCTGATGTTGTCATCTGCAATG 714
 Db 156115 GGCATCAACCTGATCGAGAGCATGACGAGCACTTCTGATGTTGTCATCTGCAACG 156056
 QY 715 CACGAGAGAGGTCATGAGCGGAGCTACAGAGCGACCTTTTGTGAGCGAGGACAC 774
 Db 156055 CACGAGAGAGGTCATGAGCGGAGCTACAGAGCGACCTTTTGTGAGCGAGGACAC 155996
 QY 775 CTTGGGATGCTGACCTTGGCAACGCTCAATTTGAGTCCAGACAGACCCGACTGA 834
 Db 155995 CTTGGGATGCTGACCTTGGGCAACGCTCAATTTGAGTCCAGACAGACCCGACTGA 155936

RESULT 15
 AC118474 192336 bp DNA linear HTG 25-MAY-2003
 AC118474 Mus musculus clone RP23-288012, *** SEQUENCING IN PROGRESS ***, 4
 LOCUS unordered pieces.
 DEFINITION AC118474
 AC118474 GI:31074905
 ACCESSION HTG; HTGS_PHASE1; HTGS_FUILLTOP; HTGS_ACTIVEFIN.
 VERSION HTG; HTGS_PHASE1; HTGS_FUILLTOP; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 192336)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP23-288012
 JOURNAL Unpublished
 AUTHORS (bases 1 to 192336)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazates, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trisillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (17-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 192336)

REFERENCE
 JOURNAL
 AUTHORS
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corn, B., Dearellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hago, B., Hago, B., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (25-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 25, 2003 this sequence version replaced gi:28975887.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

COMMENT
 TITLE
 JOURNAL
 AUTHORS
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence submissions@genome.wi.mit.edu
 Project Information
 Center Project name: L23680
 Center Clone name: 288_O_12

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 36631: contig of 36631 bp in length
 36632 36731: gap of 100 bp
 36732 160664: contig of 123933 bp in length
 160665 160764: gap of 100 bp
 160765 177955: contig of 17191 bp in length
 177956 178055: gap of 100 bp
 178056 192336: contig of 14281 bp in length.
 Location/Qualifiers
 1. 192336

```
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-288012"
/clone_11b="RPECI-23 Female Mouse BAC"
BASE COUNT      58391 a 36846 c 37829 g 58262 t 1008 others
ORIGIN
```

```
Query Match      67.9%; Score 566; DB 2; Length 192336;
Best Local Similarity 81.6%; Pred. No. 8.9e-114;
Matches 684; Conservative 0; Mismatches 140; Indels 14; Gaps 2;

QY      1  ATGGCTTCGGGGGAGAGCGGAAGACTCCGAGAGCCCGC-----GCCGAGGCCGACCGTCC 56
DB      77706 ATGGCTTCGGGGGAGAGCGGGGCGTGCACAGAGTACTGCGCCTGCCGAGGCCGATCTTCC 77765
QY      57  GCACCAGCGGCCCTTCTGATAGGGGTAGCGGCGGCACTGCCAGCGGGAAATCGACCGT 116
DB      77766 TCAGCCTTGGCCGTTCTCATCGCGGTAGTGCAGCGCACCGCTAGTGGCAAGTCAACAGT 77825
QY      117 GTGTGAGAAGATCATGAGGTGCTGGGACAGACGAGGTGAGACAGCGGACGGGAAGT 176
DB      77826 GTGTGAGAAGATCATGAGGTGCTGGGACAGATGAGTGAAGTCCGCTAGCGCAAGTT 77885
QY      177 GGTGATCTTGAGCCAGAGCAGGTTCTACAGGTCCTGACGCGCAGAGCAGAAAGCCAAAGC 236
DB      77886 GGTGATCTTGAGCCAGAGCAGGTTCTACAGGTTCTGACGCGTGAAGAGCCAAAGC 77945
QY      237 CTTGAAAGGACAGTACATTTTGGACATCCAGATGCTTTGATATGATTTGATGCACAG 296
DB      77946 TTTGAAAGGACAGTACATTTTGGACATCCAGATGCTTTGATATGATCTGATGACAA 78005
QY      297 GACTCTGAGAAGACATCGTGAAGGGCAAAAGGTGAGGTGCGGACCTATGATTTTGTGAC 356
DB      78006 GACTCTGAGAAGACATTTATTTGAAGGCAAAAGTTTGAAGTCTTATGATTTTGTGAC 78065
QY      357 AACATCAAGGTTTACAGAGACCAAGTGTCTACCTGCGGACGTGTTCTGTTGAGGG 416
DB      78066 CCACCTCAAGGTTTACAGAGACCAAGTGTCTACCAATGATGCGGTTCTGTTGAGGG 78125
QY      417 CATCTTGCTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTGCGGCTCTTCTGTGA 476
DB      78126 CATCTTGCTGTTCTACACTCAAGAGATCCGGGACATGTTTCACTGCGGCTCTTGTGA 78185
QY      477 CACCGATCCGAGCGTCAAGCTGTCTCGAAGATTCTCCGGGACGTGCGCGAGGAGGGA 536
DB      78186 CACGATCCCTGATGTTAGGCTGTCTTGAAGATTCTCGGATGTGCATGAGGAGGGA 78245
QY      537 CCTGAGCAGATTTGAGCAGTACACCACTTCGTGAAGCGGCTTCGAGAGTTCTG 596
DB      78246 CCTGAGCAGATTCCTGACTGAGTCAACCACTTGTGAATCGGCTTCGAGAGTTCTG 78305
QY      597 CCTGCCGCAAAAGTATGCCGATGTGATCATCCACGAGAGTGAACAATATGTTGC 656
DB      78306 CCTGCCGCTTAGA-----GTGATCATCCCTCGAGGCTTGACATATGTGTC 78355
QY      657 CATCAACCTGATCGAGCAGATCCAGAGCAATTCTGAATGTGACATCTGCAATAGGCA 716
DB      78356 CATCAACCTGATCGATTAACATCCAGAGCATCTTCAACGGGACCTGTGCAAGCGCA 78415
QY      717 CCGAGAGAGTCAATGAGCGAGCTACAGCGGACCTTTTCTGAGCCAGGGGACCAACC 776
DB      78416 CCGAGAGAGGCCCAATGAGCGACAGCCACAGAGGACCTTACCGAGTCAAGAGATCAACC 78475
QY      777 TGGGATGCTGACCTCTGCAAAAGGTCACATTTGAGTCCAGCAGACAGACCCCACTGA 834
DB      78476 TGGGATGCTGCACTGTGAACGCTGACACTGAGTCTAGCAGCAGATCCCATTTGA 78533
```

Search completed: November 25, 2003, 00:22:05
Job time : 3247.66 secs

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Auandi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR P-PSDB; AAM39502.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 861; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
 Query Match 100.0%; Score 834; DB 22; Length 1288;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCCGACCGTCCGCAC 60
 DB 95 ATGGCTTCGGGGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCCGACCGTCCGCAC 154
 QY 61 CAGGGGCCCCCTCCCTGATAGGGGTGAGCGGGCGGCACTGCCAGCGGAAGTGCAGTGTGT 120
 DB 155 CAGGGGCCCCCTCCCTGATAGGGGTGAGCGGGCGGCACTGCCAGCGGAAGTGCAGTGTGT 214
 QY 121 GAGAAATCATGATGAGTGTCTGGACAGAACGAGGTGMAACGGCGGACGGGAGGTGTC 180
 DB 215 GAGAAATCATGATGAGTGTCTGGACAGAACGAGGTGMAACGGCGGACGGGAGGTGTC 274
 QY 181 ATCTGAGCCAGACGAGGTCTTCAAGGTCTTGAACGCGACAGACGAGGCGCAAGGCTTGG 240
 DB 275 ATCTGAGCCAGACGAGGTCTTCAAGGTCTTGAACGCGACAGACGAGGCGCAAGGCTTGG 334
 QY 241 AAGAGCAATGATATTTTGAACATCCAGATGCTTTGATATATATTTGATGACAGACT 300
 DB 335 AAGAGCAATGATATTTTGAACATCCAGATGCTTTGATATATATTTGATGACAGACT 394
 QY 301 CTGAAGAAATCGTGAAGGGGCAAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 360
 DB 395 CTGAAGAAATCGTGAAGGGGCAAAACGATGAGGTGCGACCTATGATTTTGTGACACAC 454
 QY 361 TCAAGGTTACAGAGACACCGTGTCTACCTTCGAGCGTGTCTTGTGTTGAGGCGATC 420
 DB 455 TCAAGGTTACAGAGACACCGTGTGTCTACCTTCGAGCGTGTCTTGTGTTGAGGCGATC 514
 QY 421 TTGGTGTTCATACACCGAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTCGACACC 480
 DB 515 TTGGTGTTCATACACCGAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTCGACACC 574
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCGGGACGTCGAGCGGAGGAGACTG 540

DB 575 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGAGACTGCGCGGAGGAGGACCTG 634
 QY 541 GAGCAGATTCGTGACGAGATACACACCTTCGTGAAGCGGCTTCGAGGAGTTCTGCTTG 600
 DB 635 GAGCAGATTCGTGACGAGATACACACCTTCGTGAAGCGGCTTCGAGGAGTTCTGCTTG 694
 QY 601 CCGACAAAGAGATGATGCTGATGATCATTCCACGAGAGTGGACAAATATGTTGGCATC 660
 DB 695 CCGACAAAGAGATGATGCTGATGATCATTCCACGAGAGTGGACAAATATGTTGGCATC 754
 QY 661 AACCTGATCGTGAAGACATCCAGACATTCGTAATGATGACATCTGCAATATGACACCGA 720
 DB 755 AACCTGATCGTGAAGACATCCAGACATTCGTAATGATGACATCTGCAATATGACACCGA 814
 QY 721 GAGGGTCCAAATGAGGCGGAGCTCAAGCGGACCTTTTCTGAGCGGAGGACCACTGGG 780
 DB 815 GAGGGTCCAAATGAGGCGGAGCTCAAGCGGACCTTTTCTGAGCGGAGGACCACTGGG 874
 QY 781 ATGCTGACCTTCGGCAAGGTCATCTTGGATCCAGACGACACCCACTGA 834
 DB 875 ATGCTGACCTTCGGCAAGGTCATCTTGGATCCAGACGACACCCACTGA 928
 RESULT 2
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 AC AAD27186;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 KW Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticonvulsant; ds.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FH 94..927
 FT CDS /tag= a
 FT /product= "Human 57658 protein"
 FT misc_feature 94..924
 FT /tag= b
 FT /note= "This region is specifically claimed as
 SEQ ID NO: 3 in claim 1 of the specification"
 XX
 PD WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 DR WPI; 2002-140991/18.
 DR P-PSDB; AAI6592.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX

PS Claim 1; Fig 1a; 103bp; English.

XX The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as haematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.

XX Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Query Match 100.0%; Score 834; DB 24; Length 1624;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGGAGCGAAGACTGCGAGGCCCGCGCGGAGCCGCTCCGAC 60
DB ATGGCTTCGGCGGGAGCGAAGACTGCGAGGCCCGCGCGGAGCCGCTCCGAC 153
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAGTGAACCTGTGT 120
DB 154 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGGGAAGTGAACCTGTGT 213
QY 121 GAGAGATCATGAGAGTTGCTGGACAGAACGAGGTGAGCAGCGGCGAAGGTGTC 180
DB 214 GAGAGATCATGAGAGTTGCTGGACAGAACGAGGTGAGCAGCGGCGAAGGTGTC 273
QY 181 ATCTGAGCCGAGGACAGGTCTCAAGGTCCTGACCGCAGAGCAAGAGGCGCTTG 240
DB 274 ATCTGAGCCGAGGACAGGTCTCAAGGTCCTGACCGCAGAGCAAGAGGCGCTTG 333
QY 241 AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATATGATGACAGAGACT 300
DB 334 AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATATGATGACAGAGACT 333
QY 301 CTGAGAGACATCGTGGAGGGCAAAACGGTGGAGGTGCGACTATGATTTTGTGACAC 360
DB 394 CTGAGAGACATCGTGGAGGGCAAAACGGTGGAGGTGCGACTATGATTTTGTGACAC 453
QY 361 TCAAGGTTCCAGAGACACCGGTGCTACCCCTGCGGACGTGTTCTGTTGAGGGCTC 420
DB 454 TCAAGGTTCCAGAGACACCGGTGCTACCCCTGCGGACGTGTTCTGTTGAGGGCTC 513
QY 421 TTGCTTTTACAGCAGGAGATCCGGGACATGTTCCACCTGCGGCTCTTCTGTGACAC 480
DB 514 TTGCTTTTACAGCAGGAGATCCGGGACATGTTCCACCTGCGGCTCTTCTGTGACAC 573
QY 481 GACTCCGAGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGAGGAGGAGCTG 540
DB 574 GACTCCGAGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGAGGAGGAGCTG 633
QY 541 GAGCAGATTCAGCGAGATCAACCACTTCGTGTAACCCGGCTTCGAGAGATTCTGCTG 600
DB 634 GAGCAGATTCAGCGAGATCAACCACTTCGTGTAACCCGGCTTCGAGAGATTCTGCTG 693
QY 601 CCGACAAGAGATATCCGATGATCATCCACGAGAGTGGACATATAGTTGTCATC 660
DB 694 CCGACAAGAGATATCCGATGATCATCCACGAGAGTGGACATATAGTTGTCATC 753
QY 661 AACCTGATGTGACAGACATCCAGACATTTGTAATGTGACATCTGCAATGGACCGA 720
DB 754 AACCTGATGTGACAGACATCCAGACATTTGTAATGTGACATCTGCAATGGACCGA 813

QY 721 GGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGCACCTGGG 780
DB 814 GGAGGTCCTCAATGGGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGCACCTGGG 873
QY 781 ATGCTGACCTCTGGCAACGCTCACATTTTGAGTCCAGACAGACCCCACTGA 834
DB 874 ATGCTGACCTCTGGCAACGCTCACATTTTGAGTCCAGACAGACCCCACTGA 927

RESULT 3

AAH15621 standard; cDNA; 2160 BP.

XX AAH15621;
XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; se.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000BP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JUN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 13951; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises at least 15 nucleotides, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination

CC of the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

Query Match 100.0%; Score 834; DB 22; Length 2160;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 ATGGCTTCGGCGGAGGCGAAGACTGAGAGAGCCCGCGGAGGCGGACCCGCGCAC 60
Db 95 ATGGCTTCGGCGGAGGCGAAGACTGAGAGAGCCCGCGGAGGCGGACCCGCGCAC 154
Qy 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCGACTCCAGCGGGAAGTCGACCGTGT 120
Db 155 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCGACTCCAGCGGGAAGTCGACCGTGT 214
Qy 121 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 180
Db 215 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTCTC 274
Qy 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGGCAGAGGAGGCGGACCTTG 240
Db 275 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTGACGGCAGAGGAGGCGGACCTTG 334
Qy 241 AAGAGCAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
Db 335 AAGAGCAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGATGACAGGACT 394
Qy 301 CTGAGAGACATCTGAGGCGAAGGCGTGGAGGTGCGGACTTATGATTTTGTGACAC 360
Db 395 CTGAGAGACATCTGAGGCGAAGGCGTGGAGGTGCGGACTTATGATTTTGTGACAC 454
Qy 361 TCAGAGTTACAGAGACACGCGTGTCTACCCCTGCGGACGCGTCTGTTGAGGGATC 420
Db 455 TCAGAGTTACAGAGACACGCGTGTCTACCCCTGCGGACGCGTCTGTTGAGGGATC 514
Qy 421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTCTTCTGTGACAC 480
Db 515 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTCTTCTGTGACAC 574
Qy 481 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGGAGGAGGACTTG 540
Db 575 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGACGTGCGCGGAGGAGGACTTG 634
Qy 541 GAGAGATTCAGAGGACGAGTACACCACTTGTGAGGCGGCTTTCGAGGAGTCTGCTG 600
Db 635 GAGAGATTCAGAGGACGAGTACACCACTTGTGAGGCGGCTTTCGAGGAGTCTGCTG 694
Qy 601 CCGCAGAAAGATATGCGCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
Db 695 CCGCAGAAAGATATGCGCATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 754
Qy 661 AACCTGATCGTGACGACATCCAGAGCAATTTGATGATGATCTGCAATATGCGACCGA 720
Db 755 AACCTGATCGTGACGACATCCAGAGCAATTTGATGATGATCTGCAATATGCGACCGA 814
Qy 721 GAGAGGTCATGAGGCGGAGCTACAGCGACCTTTCTGAGCGAGGGGACCCCTGGG 780
Db 815 GAGAGGTCATGAGGCGGAGCTACAGCGACCTTTCTGAGCGAGGGGACCCCTGGG 874
Qy 781 ATGTGACCTCTGCGAAGCGTCAATTTGAGTTCAGAGCAGAGACCCCACTGA 834
Db 875 ATGTGACCTCTGCGAAGCGTCAATTTGAGTTCAGAGCAGAGACCCCACTGA 928

```

RESULT 4
ID AAK98735
AAK98735 standard; DNA; 834 BP.

XX AAK98735;

XX 02-MAY-2002 (first entry)

XX DNA of a human uridine kinase (UDK).

DE

KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localisation study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1..780
XX /*tag= a
XX /partial
XX /note= "No stop codon"

PN W0200172963-A2.

PD 04-OCT-2001.

PF 27-MAR-2001; 2001WO-US09663.

PR 27-MAR-2000; 2000US-0536647.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Ho YS, Johnson RK;

PS WPI; 2001-626259/72.

DR P-PSDB; AAO14412.

PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -

PS Claim 5; Page 22-23; 31pp; English.

XX The invention relates to newly identified human uridine kinase (UDK)
XX polypeptides and polynucleotides and methods for producing such
XX polypeptides by recombinant techniques. Also disclosed in the invention
XX are methods for utilizing uridine kinase polypeptides and polynucleotides
XX in diagnostic assays. The polynucleotides and polypeptides of the
XX invention may be used as diagnostic reagents by detecting mutations in an
XX associated gene. An array of oligonucleotide probes comprising the
XX uridine kinase polynucleotide sequence or fragments thereof can be
XX constructed to conduct efficient screening of genetic mutations, for
XX example. Detection of abnormally decreased or increased levels of
XX polypeptide or mRNA expression may also be used for diagnosing or
XX determining susceptibility of a subject to a disease of the invention.
XX The polynucleotide sequences of the invention can be used for chromosome
XX localisation studies and tissue expression studies. The polypeptides of
XX the invention or fragments thereof may be used as immunogens to produce
XX antibodies. These antibodies may be employed to isolate or identify
XX clones expressing the polypeptide. The polypeptides and polynucleotides
XX of the invention can be used as a vaccine or in gene therapy to treat
XX diseases such as human ovarian cancer, human colon carcinomas, and
XX immunological disorders. This polynucleotide sequence represents the DNA
XX of a human uridine kinase of the invention.

SQ Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 93.9%; Score 783; DB 23; Length 834;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 60
Db 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 60
Qy 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGACTCCAGCGGGAAGTCGACGTTGT 120
Db 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGACTCCAGCGGGAAGTCGACGTTGT 120
Qy 121 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACAGCGGACGCGGAAGTGTCTC 180
Db 121 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACAGCGGACGCGGAAGTGTCTC 180

```


QY 181 ATCTGAGCCAGAGCAGGTTCTAACAGTCTGACGGCAGAGCAGAGGCCAAGGCGCTTG 240
 Db 181 ATCTGAGCCAGAGCAGGTTCTAACAGTCTGACGGCAGAGGCCAAGGCGCGCTTG 240
 QY 241 AAGGACAGTACATTTTGAACATCCAGATCCCTTGTATATGATTTGAATCAAGAGACT 300
 Db 241 AAGGACAGTACATTTTGAACATCCAGATCCCTTGTATATGATTTGAATCAAGAGACT 300
 QY 301 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGGACACAC 360
 Db 301 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGGACACAC 360
 QY 361 TCAAGGTTACAGAGACACAGGATCCGAGATGTTCCACCTGCGCCTCTTCTGAGACACC 420
 Db 361 TCAAGGTTACAGAGACACAGGATGTTCCACCTGCGCCTCTTCTGAGACACC 420
 QY 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCCTCTTCTGAGACACC 480
 Db 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCCTCTTCTGAGACACC 480
 QY 481 GACTCCGACGTCAGGCTGTCTCCAGAGTTCTCCGGACGTCGCGGAGGAGGAGACTTG 540
 Db 481 GACTCCGACGTCAGGCTGTCTCCAGAGTTCTCCGGACGTCGCGGAGGAGGAGACTTG 540
 QY 541 GAGCAGATTCTGACGCGAGTACACCACTTGTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 600
 Db 541 GAGCAGATTCTGACGCGAGTACACCACTTGTGTAAGCGGCGCTTCCAGAGAGTTCTGCTG 600
 QY 601 CCGACAAAGAGATGCGCATGTGATCATCCAGAGAGTGGACATATGTTGCCATC 660
 Db 601 CCGACAAAGAGATGCGCATGTGATCATCCAGAGAGTGGACATATGTTGCCATC 660
 QY 661 AACCTGATCGTGACGACATCCAGAGATCTGTAATGTAATCTGCAAAATGACACCGA 720
 Db 661 AACCTGATCGTGACGACATCCAGAGATCTGTAATGTAATCTGCAAAATGACACCGA 720
 QY 721 GAGAGGTCATGAGGCGAGCTACAGCGGACCTTTCTAGGCGAGGAGCAACCTGGG 780
 Db 721 GAGAGGTCATGAGGCGAGCTACAGCGGACCTTTCTAGGCGAGGAGCAACCTGGG 780
 QY 781 ATCTGACCTCTGCGCAACCGGTCAATTTGAGTCCAGAGACCCCACTGA 834
 Db 781 ATCTGACCTCTGCGCAACCGGTCAATTTGAGTCCAGAGACCCCACTGA 834
 RESULT 5
 AA160444
 ID AA160444 standard; cDNA; 2152 BP.
 XX
 AC AA160444;
 XX
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 4433.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KM leukemia; ss.
 XX
 OS Homo sapiens.
 PN WO200153112-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-048725.
 PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSER INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB: AAM41288.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 PS
 PS Claim 1; SEQ ID NO 4433; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;
 Query Match 93.9%; Score 783; DB 22; Length 2152;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCGTCTGCGCGGAGCGCAAGACTGCGAGACCCCGCGAGGCGCAAGCTCCGCAC 60
 Db 59 ATGCGTCTGCGCGGAGCGCAAGACTGCGAGACCCCGCGAGGCGCAAGCTCCGCAC 118
 QY 61 CAGCGGCGCTTCTGATGAGGCGTGAAGCGGCGACCTGCGAGGAGTGCACCGTGT 120
 Db 119 CAGCGGCGCTTCTGATGAGGCGTGAAGCGGCGACCTGCGAGGAGTGCACCGTGT 178
 QY 121 GAGAGATCATGAGTGTGCGGAGCAAGAGGTGGAACAGCGGAGCGGAAGGTGTC 180
 Db 179 GAGAGATCATGAGTGTGCGGAGCAAGAGGTGGAACAGCGGAGCGGAAGGTGTC 238
 QY 181 ATCTGAGCCAGAGCAGGTTCTTCAAGGCTCTGACGCGCAGAGCAAGAGCCCAAGCCTTG 240
 Db 239 ATCTGAGCCAGAGCAGGTTCTTCAAGGCTCTGACGCGCAGAGCAAGAGCCCAAGCCTTG 298
 QY 241 AAGGACAGTACATTTTGAACATCCAGATCCCTTGTATATGATTTGAATCAAGAGACT 300
 Db 299 AAGGACAGTACATTTTGAACATCCAGATCCCTTGTATATGATTTGAATCAAGAGACT 358
 QY 301 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGGACACAC 360
 Db 359 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGGACACAC 418
 QY 361 TCAAGGTTACAGAGACACAGGTTCTTACCTGCGGAGAGTGTCTGTTTGAAGGCAATC 420
 Db 419 TCAAGGTTACAGAGACACAGGTTCTTACCTGCGGAGAGTGTCTGTTTGAAGGCAATC 478
 QY 421 TTGGTGTCTACAGCCAGAGATCCGAGATGTTCCACCTGCGCCTCTTCTGAGACACC 480

```

Db      479  TTGGGTTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGTGGACACC 538
Qy      481  GACTCCGACCTCAGGCTCTCTCGAAGAGTTTCCGGGACGTCGCCGAGGAGGACCTG 540
Db      539  GACTCCGACCTCAGGCTCTCTCGAAGAGTTTCCGGGACGTCGCCGAGGAGGACCTG 598
Qy      541  GAGCAGATTCTGACGACATACCACTTCGTAAGCCGGCTTCGAGAGATTCTGACCTG 600
Db      599  GAGCAGATTCTGACGACATACCACTTCGTAAGCCGGCTTCGAGAGATTCTGACCTG 658
Qy      601  CCGACAAAGAGTATGCGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 660
Db      659  CCGACAAAGAGTATGCGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 718
Qy      661  AACTGATTCGTCGACGACATCCAGACATTTCTGAATGTAATGTAATGTAATGTAAT 720
Db      719  AACTGATTCGTCGACGACATCCAGACATTTCTGAATGTAATGTAATGTAATGTAAT 778
Qy      721  GGAGGGTTCATGAGCGGAGCTACAGCGGACCTTTTCTGAGCCGAGGAGCCACCTGG 780
Db      779  GGAGGGTTCATGAGCGGAGCTACAGCGGACCTTTTCTGAGCCGAGGAGCCACCTGG 838
Qy      781  ATGCTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCCACTGA 834
Db      839  ATGCTGACCTCTGCGAAACGGTCACTTTGAGTTCAGCAGCAGACCCCACTGA 892

```

RESULT 6
AAH04832
ID AAH04832 standard; cDNA; 753 BP.

AAH04832;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1667.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602

full-length cDNAs defined in the specification, and for the detection

and/or diagnosis of the abnormality of the proteins encoded by the

full-length cDNAs -

Claim 1; SEQ ID 1667; 2537bp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602

full-length cDNAs defined in the specification. Where a primer set

comprises: (a) an oligo-dT primer and an oligonucleotide complementary

to the complementary strand of a polynucleotide which comprises one of

the 5602 nucleotide sequences defined in the specification, where the

oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;

Query Match 68.9%; Score 575; DB 22; Length 753;

Best Local Similarity 100.0%; Pred. No. 1.6e-289; Mismatches 0; Indels 0; Gaps 0;

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1  ATGGCTTCGCGGAGGCGAAGACTGCGAGGCGCGCGCGCGGAGGCGGACCTCCGAC 60
Db      95  ATGGCTTCGCGGAGGCGAAGACTGCGAGGCGCGCGCGCGGAGGCGGACCTCCGAC 154
Qy      61  CAGCGGCTTCTGATGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 120
Db      155  CAGCGGCTTCTGATGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 214
Qy      121  GAGAAGATCATGAGTTCGCGGAGCAAGAGAGTGAACAGGCGGAGGAGGAGGAGT 180
Db      215  GAGAAGATCATGAGTTCGCGGAGCAAGAGAGTGAACAGGCGGAGGAGGAGGAGT 274
Qy      215  GAGAAGATCATGAGTTCGCGGAGCAAGAGAGTGAACAGGCGGAGGAGGAGGAGT 274
Db      181  ATCCTGAGCCAGGACAGTTCCTCAAGGTCCTGACGCGCAGAGCAGAGGCGGAGG 240
Qy      275  ATCCTGAGCCAGGACAGTTCCTCAAGGTCCTGACGCGCAGAGCAGAGGCGGAGG 334
Db      241  AAAGACAGTACATTTTGAACATCCAGATGCTTTGATGATGATTTGATGACAGGACT 300
Qy      335  AAAGACAGTACATTTTGAACATCCAGATGCTTTGATGATGATTTGATGACAGGACT 394
Db      301  CTGAAGAACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
Qy      395  CTGAAGAACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 454
Db      361  TCAAGGTTACAGAGACACAGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 420
Qy      455  TCAAGGTTACAGAGACACAGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 514
Db      421  TTGGTGTTCACAGCAGAGATCCGGAGATGTTCCACCTGCGCTCTTCTGTAACACC 480
Qy      515  TTGGTGTTCACAGCAGAGATCCGGAGATGTTCCACCTGCGCTCTTCTGTAACACC 574
Db      481  GACTCGAGGTCGAGGTCTCGAAGATTCCTCGGAGAGTGGCGGAGGAGGAGGAGCTG 540
Qy      575  GACTCGAGGTCGAGGTCTCGAAGATTCCTCGGAGAGTGGCGGAGGAGGAGGAGCTG 634
Db      541  GAGCAGATTCTGACGACATACCACTTCGTAAGCCGGCTTCGAGAGATTCTGACCTG 575
Qy      635  GAGCAGATTCTGACGACATACCACTTCGTAAGCCGGCTTCGAGAGATTCTGACCTG 669

```

RESULT 7
AAH75355
ID AAH75355 standard; cDNA; 900 BP.

AAH75355;

02-OCT-2001 (first entry)

DE Human uridine kinase encoding cDNA.
 XX Human; uridine kinase; UK; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 8..838
 FT /product= "uridine kinase"
 FT /note= "Claimed in claim 1"
 FT
 XX CN1287172-A.
 XX 14-MAR-2001.
 XX PD
 XX PF 07-SEP-1999; 99CN-0118818.
 XX PR 07-SEP-1999; 99CN-0118818.
 XX PA (UYFU-) UNIV FUDAN.
 XX PI Yu L, Zhao Y, Zhang H;
 XX WPI; 2001-409529/44.
 DR P-PSDB; AAG64506.
 XX
 XX Human uridine kinase and its coding sequence, preparation and
 PT application -
 PS Claim 1; Page 15(Disclosure); 20pp; Chinese.
 CC The invention relates to human uridine kinase (UK).
 CC
 CC Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
 SQ
 Query Match 62.0%; Score 517; DB 22; Length 900;
 Best Local Similarity 99.8%; Pred. No. 3e-259;
 Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 172 AAGGTGTCATCTCTGAGCCAGAGAGGTTCTACAGGTCTCTGACGAGAGAGAGAGGCC 231
 DB 176 AAGGTGTCATCTCTGAGCCAGAGAGGTTCTACAGGTCTCTGACGAGAGAGAGAGGCC 235
 QY 232 AAGGCTTGAAGAGACGTAACATTTGACCATTCAGATGCTTGTATATGATTGATG 231
 DB 236 AAGGCTTGAAGAGACGTAACATTTGACCATTCAGATGCTTGTATATGATTGATG 235
 QY 292 CACAGACTCTGAAGAATCGTGGAGGCAAAACGGTGGAGGTCGACCTATGATTTT 351
 DB 296 CACAGACTCTGAAGAATCGTGGAGGCAAAACGGTGGAGGTCGACCTATGATTTT 355
 QY 352 GTGACACTCAGAGTTACAGAGACCAAGTGTCTACCTGGCGACGTGTTCTGTTT 411
 DB 356 GTGACACTCAGAGTTACAGAGACCAAGTGTGTACCTGGCGACGTGTTCTGTTT 415
 QY 412 GAGGCACTTTGTTGTTCTACAGCCAGAGATCCGGACATGTTCCACCTGCCCTTTC 471
 DB 416 GAGGCACTTTGTTGTTCTACAGCCAGAGATCCGGACATGTTCCACCTGCCCTTTC 475
 QY 472 GTGACACCGACTCCGAGTCGAGGCTGTCTGAAGAGTTCTCCGGGACGTGGCCGAGG 531
 DB 476 GTGACACCGACTCCGAGTCGAGGCTGTCTGAAGAGTTCTCCGGGACGTGGCCGAGG 535
 QY 532 AGGACCTGGAGAGATTTGAGCCAGTACCAACCTTCTGTAAGCCGGCTTGAAGAG 591
 DB 536 AGGACCTGGAGAGATTTGAGCCAGTACCAACCTTCTGTAAGCCGGCTTGAAGAG 595
 QY 592 TTCTGCTGCGGCAAGAGATATCGCATGTATATCCACAGAGAGTGGACATATG 651
 DB 596 TTCTGCTGCGGCAAGAGATATCGCATGTATATCCACAGAGAGTGGACATATG 655
 QY 652 GTTGCCATCAACTGATGTGACGACATCCAGGACATCTGATGATGATCTGCAAA 711

DB 656 GTTGCCATCAACTGATGTGACGACATCCAGGACATTTGATGTGACATCTGCAAA 715
 QY 712 TGGCACCAGAGAGGTCATGAGCCGA 739
 DB 716 TGGCACCAGAGAGGTCATGAGCCGA 743
 RESULT 8
 ID ABL89762 standard; cDNA; 1396 BP.
 AC ABL89762;
 XX
 XX 24-MAY-2002 (first entry)
 DE Human polynucleotide SEQ ID NO 324.
 XX
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200190304-A2.
 XX
 XX 29-NOV-2001.
 XX
 XX 18-MAY-2001; 2001WO-US16450.
 XX
 XX 19-MAY-2000; 2000US-205515P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 PI WPI; 2002-122018/16.
 DR P-PSDB; ABB89353.
 DR
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.
 XX
 XX The invention relates to novel genes (AB189449-AB190853) and proteins
 CC (AB189449-AB190444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer; e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;
 Query Match 62.0%; Score 517; DB 24; Length 1396;
 Best Local Similarity 99.6%; Pred. No. 3e-259;
 Matches 737; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGCTTCGGCGGAGGAGAAAGTCTGAGAGACCCCGCGCGAGAGCCGACCGTCCGAC 60
 Db 41 ATGGCTTGGCGGAGGAGGAGAAAGTCTGAGAGACCCCGCGCGAGAGCCGACCGTCCGAC 100
 QY 61 CAGCGGCGCTTCTGATAGGGGTGAGCGCGGACCTG⁵CAGCGGGAAAGTCAACCGTGTGT 120
 Db 101 CAGCGGCGCTTCTGATAGGGGTGAGCGCGGACCTG⁵CAGCGGGAAAGTCAACCGTGTGT 160
 QY 121 GAGAAATCATGGAATTTCTGGGACAGAAACGAGGTGGAACAGCGGACACCGAAGTGTCT 180
 Db 161 GAGAAATCATGGAATTTCTGGGACAGAAACGAGGTGGAACAGCGGACACCGAAGTGTCT 220
 QY 181 ATCTGAGCCAGAGCAGGTTCTACAAAGTCTCTGACGAGAGACAGAAAGCCAGGCTTGT 240
 Db 221 ATCTGAGCCAGAGCAGGTTCTACAAAGTCTCTGACGAGAGACAGAAAGCCAGGCTTGT 280
 QY 241 AAAGGACAGTCAATTTTGAACCATCAAGATGCTTTGATTAATGATTTGATGACAGACT 300
 Db 281 AAAGGACAGTCAATTTTGAACCATCAAGATGCTTTGATTAATGATTTGATGACAGACT 340
 QY 301 CTGAAGAACATCTGAGAGGAGGAGAAACGGTGGAGGTCCGACCTATGATTTTGTACACAC 360
 Db 341 CTGAAGAACATCTGAGAGGAGGAGAAACGGTGGAGGTCCGACCTATGATTTTGTACACAC 400
 QY 361 TCAAGTTACAGAGACACGAGTGTCTACCTCGGAGAGTGTCTGTTTGAAGGACATC 420
 Db 401 TCAAGTTACAGAGACACGAGTGTCTACCTCGGAGAGTGTCTGTTTGAAGGACATC 460
 QY 421 TTGGTGTCTACAGCAGAGAGATCCGGAGACATGTTTCACTGCGCTTCTGTGAGACAC 480
 Db 461 TTGGTGTCTACAGCAGAGAGATCCGGAGACATGTTTCACTGCGCTTCTGTGAGACAC 520
 QY 481 GACTCCGAGCGTCCGCTGTCTGAAAGATTTCTCCGGGACGTGCGCCGAGGAGGAGACCTG 540
 Db 521 GACTCCGAGCGTCCGCTGTCTGAAAGATTTCTCCGGGACGTGCGCCGAGGAGGAGACCTG 580
 QY 541 GAGCAGATTTCTGACGACATACCACTTCTGTGAAAGCGGCGCTTCAAGAGATTTCTGCTG 600
 Db 581 GAGCAGATTTCTGACGACATACCACTTCTGTGAAAGCGGCGCTTCAAGAGATTTCTGCTG 639
 QY 601 CCGACAAAGAGTATGCCAGATGTGATCTCCACGAGAGAGTGAACAATATGTTGCCATC 660
 Db 640 CCGACAAAGAGTATGCCAGATGTGATCTCCACGAGAGAGTGAACAATATGTTGCCATC 699
 QY 661 AACCGGATCGTGCAGACATCCAGAGACATCTGAAATGTTGACATCTGAAATGACACCGA 720
 Db 700 AACCGGATCGTGCAGACATCCAGAGACATCTGAAATGTTGACATCTGAAATGACACCGA 759
 QY 721 GGAGGTTCCAAATGGGCGGAG 740
 Db 760 GGAGGTTCCAAATGGGCGGAG 779
 RESULT 9
 ABSS1801
 ID ABSS1801 standard; cDNA; 734 BP.
 XX
 AC ABSS1801;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human mdt cDNA Incyte ID No: LI:235557.12:2001JUN12.
 XX
 KW Human; molecule for disease detection and treatment; MDDR; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiaschematic; gene therapy; gene; ss.
 OS
 XX Homo sapiens.
 XX
 PN WO200255738-A2.
 XX

PD 18-JUL-2002.
 XX
 XX 09-JAN-2002; 2002MO-US01008.
 XX
 PR 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-263252P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 FI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL,
 PI Dan TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen MJ,
 PI Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
 XX
 DR WPI: 2002-590679/63.
 DR P-PSDB: ABG70328.
 XX
 PT New disease detection and treatment molecule (MDPT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDR expression, e.g. autoimmune or
 PT inflammatory disorders -
 XX
 PS Claim 1; Page 98; 129pp; English.
 XX
 CC The present invention relates to the isolation of novel human
 CC molecules for disease detection and treatment (MDPT), and the
 CC polynucleotide sequences (mdpt) encoding them. The MDPT polypeptides
 CC may be used to screen for molecules that bind to, or are bound by the
 CC encoded polypeptides, and to develop a transcript image of a tissue or
 CC cell type. Probes comprising at least 20 nucleotides of the mdpt
 CC polynucleotide may be used to assess the toxicity of a test compound.
 CC The MDPT polypeptides and mdpt polynucleotides are useful in the
 CC diagnosis, study, prevention and treatment of diseases associated with
 CC the expression of molecules for disease detection and treatment. Such
 CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
 CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
 CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
 CC polynucleotides may also be used as molecule markers, in microarrays,
 CC and in somatic or germline gene therapy. ABSS1779-ABSS1814 encode
 CC the MDPT proteins of the invention.
 XX
 SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;
 Query Match 40.5%; Score 338; DB 24; Length 734;
 Best Local Similarity 100.0%; Pred. No. 7.3e-166;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 266 CAGATGCTTTGATTAATGATTTGATGACACAGACTCTGAAGAACATCTGAGAGGCAAAA 325
 Db 12 CAGATGCTTTGATTAATGATTTGATGACACAGACTCTGAAGAACATCTGAGAGGCAAAA 71
 QY 326 CGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCAAGCTGG 385
 Db 72 CGGTGAGAGTCCGACCTATGATTTTGTGACACACTCAAGTTACAGAGACCAAGCTGG 131
 QY 386 TCTACCTGCGAGAGTGTCTGTGAGAGGACATCTTGGTGTCTACAGCAGAGAGATCC 445
 Db 132 TCTACCTGCGAGAGTGTCTGTGAGAGGACATCTTGGTGTCTACAGCAGAGAGATCC 191
 QY 446 GGGACATGTTTCCACTGCGCTTCTTGTGAGACACGACTCCGACGTCAAGCTGTCTGAA 505
 Db 192 GGGACATGTTTCCACTGCGCTTCTTGTGAGACACGACTCCGACGTCAAGCTGTCTGAA 251
 QY 506 GAGTTCTCCGGGACGTGCGCCGAGGAGGAGACCTGAGCAGATTTGACGCAATACCA 565
 Db 252 GAGTTCTCCGGGACGTGCGCCGAGGAGGAGACCTGAGCAGATTTGACGCAATACCA 311
 QY 566 CTTTGTGAAGCGGCGCTTCAAGAGTGTCTGCTGCGG 603

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251859.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259578.
XX
PA (HUMA-) HUMAN GENOME SCL INC.
XX
P1 Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Example 2; SEQ ID NO 2804; 781bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (AAB03087-AB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment,
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischemias; (d) wound
CC healing; (e) neurological diseases such as cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
XX
Query Match 21.9%; Score 183; DB 22; Length 9732;
Best Local Similarity 100.0%; Pred. No. 5.8e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 652 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGATCATCTGCAAA 711
Db 3901 GTTGCCATCAACCTGATGTGACAGACATCCAGACATTTGAAATGATCATCTGCAAA 3960
QY 712 TGGCAACGAGAGAGGTCCAAATGGGCGAGACTCAAGCGGACCTTTTCTGACCCAGGGGAC 771
Db 3961 TGGCAACGAGAGAGGTCCAAATGGGCGAGACTCAAGCGGACCTTTTCTGACCCAGGGGAC 4020
QY 772 CACCTGGGATGTGACCTGTGGCAACGGTCAATTTGGAGTGCAGAGAGACCCAC 831
Db 4021 CACCTGGGATGTGACCTGTGGCAACGGTCAATTTGGAGTGCAGAGAGACCCAC 4080
QY 832 TGA 834
|||

Db 4081 TGA 4083
RESULT 11
ABX59427
ID ABX59427 standard; cDNA; 9732 BP.
XX
AC ABX59427;
XX
DT 26-FEB-2003 (first entry)
XX
DE cDNA encoding novel human musculoskeletal system antigen #1771.
XX
KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KW revascularisation; thrombosis; arteriosclerosis; mineral content;
KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
KW post-operative tissue repair; limb regeneration; neuronal growth;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW AIDS-related complex; chondrocyte growth; bone regeneration;
KW periodontal regeneration; tissue transport; bone graft; skin aging;
KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KW cell growth; organ transplant; cell differentiation; body height;
KW weight; hair colour; eye colour; skin; percentage of adipose tissue;
KW pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
KW depression; tendency for violence; pain; reproductive capability;
KW hormone level; endocrine level; appetite; libido; memory; stress;
KW storage capability; fat content; lipid content; protein content;
KW carbohydrate content; vitamin content; cofactor content;
KW nutritional component.
XX
OS Homo sapiens.
XX
EN US2002147140-A1.
XX
PD 10-OCT-2002.
XX
PF 17-JAN-2001; 2001US-0764877.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.

PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-128199/12.
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
XX associated polypeptides, useful for detecting disorders, e.g. cancer -
XX
XX Disclosure; SEQ ID NO 2804; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
XX sequence encoding musculoskeletal system associated polypeptides useful
XX for detecting disorders, e.g., cancer or cancer metastases, in animals
XX or humans. The nucleic acid: stimulates re-vascularisation of ischemic
XX tissues associated with conditions such as thrombosis, arteriosclerosis,
XX and other cardiovascular conditions; treats wounds due to injuries,
XX burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
XX and limb regeneration; stimulates neuronal growth; can treat and prevent
XX neuronal damage occurring in certain disorders or neurodegenerative
XX conditions, such as, Alzheimer's disease, Parkinson's disease, and
XX AIDS-related complex; stimulates chondrocyte growth, thus they can be
XX used to enhance bone and periodontal regeneration and aid in tissue
XX transports or bone grafts; prevents skin aging due to sunburn by
XX stimulating keratinocyte growth; prevents hair loss, since FGF family
XX members activate hair-forming cells and promotes melanocyte growth;
XX stimulates growth and differentiation of hematopoietic cells and bone
XX marrow cells when used in combination with other cytokines; maintains
XX organs before transplantation or for supporting cell culture of primary
XX tissues; induces tissue of mesodermal origin to differentiate in early
XX embryos; increases or decreases the differentiation or proliferation of
XX embryonic stem cells, besides, hematopoietic lineage; modulates
XX mammalian characteristics, such as, body height, weight, hair colour, eye
XX colour, skin, percentage of adipose tissue, pigmentation, size, and shape
XX (e.g., cosmetic surgery); modulates mammalian metabolism; changes
XX mammal's mental state or physical state by influencing biorhythms,
XX cardiac rhythms, depression, tendency for violence, tolerance for pain,
XX reproductive capabilities, hormonal or endocrine levels, appetite,
XX libido, memory, or stress; increases or decreases storage capabilities,
XX fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
XX or other nutritional components. This sequence encodes a novel human
XX musculoskeletal system antigen.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from the US patent office at
XX ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Query Match 21.9%; Score 183; DB 25; Length 9732;
Best Local Similarity 100.0%; Pred. No. 5.8e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

652 GTTGCCATCAACCGATCGTCAGACATCCAGACATTCGATGATGTCGATCGAAA 711
|||||

Db 3901 GTTGCCATCAACCGATCGTCAGACATCCAGACATTCGATGATGTCGATCGAAA 3960
Qy 712 TGGCACCAGAGAGGGTTCATATGGGCGGAGCTACACCGAACCTTTCTGAGCCAGGGGAC 771
Db 3961 TGGCACCAGAGAGGGTTCATATGGGCGGAGCTACACCGAACCTTTCTGAGCCAGGGGAC 4020
Qy 772 CACCCCTGGATGTGACCTCTGGCAACGGTGCACATTTGAGATCCAGACAGACCCAC 831
Db 4021 CACCCCTGGATGTGACCTCTGGCAACGGTGCACATTTGAGATCCAGACAGACCCAC 4080
Qy 832 TGA 834
Db 4081 TGA 4083

RESULT 12
AAL36440
ID AAL36440 standard; DNA; 19125 BP.
XX
XX AAL36440;
AC
XX 08-JAN-2002 (first entry)
XX
XX Human musculoskeletal system related polynucleotide SEQ ID NO 2805.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX
XX Homo sapiens.
XX
XX MO200155367-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225477.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0221182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231966.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
XX
XX Example 2; SEQ ID NO 2805; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
PS (AAB03087-AAB04109) associated with the musculoskeletal system useful
PS for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune hemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
QY
Db Query Match 21.9%; Score 183; DB 22; Length 19125;
Best Local Similarity 100.0%; Pred. No. 5; 9e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
652 GTTGCATCAACTGATCGTGCAGCAGCATTCGATGATGTCATTCGAAA 711
5997 GTTGCATCAACTGATCGTGCAGCAGCATTCGATGATGTCATTCGAAA 6056


```

QY      712  TGGCACCAGAGAGGCTCAATGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGAC   771
Db      6057  TGGCACCAGAGAGGCTCAATGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGAC   6116
QY      772  CACCTTGGAGTGTGACCTCTGGCAACGGTCACTTTGGAGTCCAGACAGACCCAC   831
Db      6117  CACCTTGGAGTGTGACCTCTGGCAACGGTCACTTTGGAGTCCAGACAGACCCAC   6176
QY      832  TGA 834
Db      6177  TGA 6179

RESULT 13
ABX59428
ID      ABX59428 standard; cDNA; 19125 BP.
AC      ABX59428;
XX      26-FEB-2003 (first entry)
DT      cDNA encoding novel human musculoskeletal system antigen #1772.
XX      DE
XX      Gene; ss; musculoskeletal system antigen; cancer; metastasis;
KM      re-vascularisation; thrombosis; arteriosclerosis; mineral content;
KM      post-operative tissue repair; limb regeneration; neuronal growth;
KM      neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM      AIDS-related complex; chondrocyte growth; bone regeneration;
KM      periodontal regeneration; tissue transport; bone graft; skin aging;
KM      keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
KM      cell growth; organ transplant; cell differentiation; body height;
KM      weight; hair colour; eye colour; skin; percentage of adipose tissue;
KM      pigmentation; cosmetic surgery; metabolism; biorthym; cardiac rhythm;
KM      depression; tendency for violence; pain; reproductive capability;
KM      hormone level; endocrine level; appetite; libido; memory; stress;
KM      storage capability; fat content; lipid content; protein content;
KM      carbohydrate content; vitamin content; cofactor content;
KM      nutritional component.
XX      Homo sapiens.
XX      OS
XX      PN      US2002147140-A1.
XX      PD      10-OCT-2002.
XX      PF      17-JAN-2001; 2001US-0764877.
XX      PR      31-JAN-2000; 2000US-179065P.
PR      04-FEB-2000; 2000US-180628P.
PR      28-JUN-2000; 2000US-214886P.
PR      07-JUL-2000; 2000US-216647P.
PR      07-JUL-2000; 2000US-216880P.
PR      11-JUL-2000; 2000US-217487P.
PR      11-JUL-2000; 2000US-217496P.
PR      14-JUL-2000; 2000US-218290P.
PR      26-JUL-2000; 2000US-220963P.
PR      14-AUG-2000; 2000US-224518P.
PR      14-AUG-2000; 2000US-224519P.
PR      14-AUG-2000; 2000US-225267P.
PR      14-AUG-2000; 2000US-225268P.
PR      14-AUG-2000; 2000US-225270P.
PR      14-AUG-2000; 2000US-225447P.
PR      14-AUG-2000; 2000US-225757P.
PR      14-AUG-2000; 2000US-225758P.
PR      22-AUG-2000; 2000US-226868P.
PR      30-AUG-2000; 2000US-228924P.
PR      01-SEP-2000; 2000US-229287P.
PR      01-SEP-2000; 2000US-229343P.
PR      01-SEP-2000; 2000US-229344P.
PR      01-SEP-2000; 2000US-229345P.

```

```

PR      05-SEP-2000; 2000US-229509P.
PR      05-SEP-2000; 2000US-229513P.
PR      08-SEP-2000; 2000US-231413P.
PR      21-SEP-2000; 2000US-234223P.
PR      21-SEP-2000; 2000US-234274P.
PR      25-SEP-2000; 2000US-234997P.
PR      27-SEP-2000; 2000US-235834P.
PR      29-SEP-2000; 2000US-236337P.
PR      29-SEP-2000; 2000US-236377P.
PR      29-SEP-2000; 2000US-236388P.
PR      29-SEP-2000; 2000US-236399P.
PR      29-SEP-2000; 2000US-236370P.
PR      02-OCT-2000; 2000US-236802P.
PR      02-OCT-2000; 2000US-237037P.
PR      02-OCT-2000; 2000US-237038P.
PR      02-OCT-2000; 2000US-237039P.
PR      02-OCT-2000; 2000US-237040P.
PR      13-OCT-2000; 2000US-237935P.
PR      20-OCT-2000; 2000US-240960P.
PR      20-OCT-2000; 2000US-241785P.
PR      20-OCT-2000; 2000US-241809P.
PR      01-NOV-2000; 2000US-244617P.
PR      17-NOV-2000; 2000US-249299P.
PR      08-DEC-2000; 2000US-251856P.
PR      08-DEC-2000; 2000US-251868P.
PR      08-DEC-2000; 2000US-251869P.

XX      (ROSE/) ROSEN C A.
XX      (RUBE/) RUBE S M.
XX      (BARA/) BARUSH S C.
XX      Rosen CA, Ruben SM, Barash SC,
XX      WPI; 2003-128199/12.
XX      Isolated nucleic acid molecules encoding musculoskeletal system
XX      associated polypeptides, useful for detecting disorders, e.g. cancer -
XX      disclosure, SEQ ID NO 2805; 321pp; English.
XX      CC
XX      CC The invention describes an isolated nucleic acid molecule comprising a
XX      CC sequence encoding musculoskeletal system associated polypeptides useful
XX      CC for detecting disorders, e.g., cancer or cancer metastases, in animals
XX      CC or humans. The nucleic acid stimulates re-vascularisation of ischaemic
XX      CC tissues associated with conditions such as thrombosis, arteriosclerosis,
XX      CC and other cardiovascular conditions; treats wounds due to injuries,
XX      CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
XX      CC and limb regeneration; stimulates neuronal growth; can treat and prevent
XX      CC neuronal damage occurring in certain disorders or neurodegenerative
XX      CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
XX      CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
XX      CC used to enhance bone and periodontal regeneration and aid in tissue
XX      CC transports or bone grafts; prevents skin aging due to sunburn by
XX      CC stimulating keratinocyte growth; prevents hair loss, since FGF family
XX      CC members activate hair-forming cells and promotes melanocyte growth;
XX      CC stimulates growth and differentiation of hematopoietic cells and bone
XX      CC marrow cells when used in combination with other cytokines; maintains
XX      CC organs before transplantation or for supporting cell culture of primary
XX      CC tissues; induces tissue of mesodermal origin to differentiate in early
XX      CC embryonic; increases or decreases the differentiation or proliferation of
XX      CC embryonic stem cells, besides, hematopoietic lineage; modulates
XX      CC mammalian characteristics, such as, body height, weight, hair colour, eye
XX      CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
XX      CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
XX      CC mammal's metal state or physical state by influencing biohythms,
XX      CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
XX      CC reproductive capabilities, hormonal or endocrine levels, appetite,
XX      CC libido, memory, or stress; increases or decreases storage capabilities,
XX      CC fat content, lipid, protein, carbohydrate, vitamin, minerals, cofactors
XX      CC or other nutritional components. This sequence encodes a novel human
XX      CC musculoskeletal system antigen.
XX      CC Note: The sequence data for this patent did not form part of the
XX      CC printed specification, but was obtained in electronic format directly

```

CC from the US patent office at
CC ftp://seqdata.uspto.gov/sequence.html?docid=20020147140.
XX
SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
Query Match 21.9%; Score 183; DB 25; Length 19125;
Best Local Similarity 100.0%; Pred. No. 5.9e-85;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 GTTGCATCAACCTGATGTGACATCCAGACATCTTGATGTGACATCTGCAAA 711
DB 5997 GTTGCATCAACCTGATGTGACATCCAGACATCTTGATGTGACATCTGCAAA 6056
QY 712 TGGCACCAGAGAGGCTCCAAATGCGCGAGCTACTAAAGCGAGCTTTCTGAGCCAGGGGAC 771
DB 6057 TGGCACCAGAGAGGCTCCAAATGCGCGAGCTACTAAAGCGAGCTTTCTGAGCCAGGGGAC 6116
QY 772 CACCTTGGGATGTGACCTTGGCAAAAGGTCACATTTGAGTCCAGCAGAGACCCGAC 831
DB 6117 CACCTTGGGATGTGACCTTGGCAAAAGGTCACATTTGAGTCCAGCAGAGACCCGAC 6176
QY 832 TGA 834
DB 6177 TGA 6179

RESULT 14
ABN51411
ID ABN51411 standard; DNA; 65 BP.
AC ABN51411;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:24159.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
FN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S,
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
PS Example 1; SEQ ID 24159; 47bp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in

CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 16 A; 13 C; 22 G; 14 T; 0 other;
Query Match 3.1%; Score 26; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 493 AGGCTGTCTGAGAGATTCCTCCGGA 518
DB 16 AGGCTGTCTGAGAGATTCCTCCGGA 41

RESULT 15
AAZ15454
ID AAZ15454 standard; cDNA; 715 BP.
AC AAZ15454;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:2923.
XX
KW Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
FN WO9938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Ctkvenjakov R, Dickson M, Dymnac R, Dymnac S;
PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Suduth-Klinger J, Williams LR;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1410-1411; 2479pp; English.
XX
CC The present invention describes a library of human polynucleotides
CC comprising the sequences given in AAZ12532 to AAZ1779. Also described is
CC a method of detecting differentially expressed genes correlated with the

CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX
SQ Sequence 715 BP, 179 A, 189 C, 178 G, 157 T, 12 other;

Query Match 2.8%; Score 23; DB 20; Length 715;

Best Local Similarity 100.0%; Pred. No. 0.17;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 667 ATCGTGAGCAGCATCCAGAGACAT 689
DB 572 ATCGTGAGCAGCATCCAGAGACAT 594

Search completed: November 25, 2003, 02:31:21
Job time : 258.154 secs

This Page Blank (uspto)

PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM39502.
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 861; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;
 XX
 Query Match 100.0%; Score 834; DB 22; Length 1288;
 Best Local Similarity 100.0%; Pred. No. 8.2e-206;
 Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCGCCGCTCCGCAC 60
 DB 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAGACCCCGCGCGGAGCGCCGCTCCGCAC 154
 QY 61 CAGCGGCCCTCTCTGATRGGGGTGAACCGCGCGCACTGCGACCGGGAAGTGCATCTGTGT 120
 DB 155 CAGCGGCCCTCTCTGATRGGGGTGAACCGCGCGCACTGCGACCGGGAAGTGCATCTGTGT 214
 QY 121 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAACGCGGACGCGGAGGTGTGTC 180
 DB 215 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAACGCGGACGCGGAGGTGTGTC 274
 QY 181 ATCTGAGCCAGACAGAGTCTCAAGGTCTCTGACGCGAGACAGAAAGCCAGGCTTGG 240
 DB 275 ATCTGAGCCAGACAGAGTCTCAAGGTCTCTGACGCGAGACAGAAAGCCAGGCTTGG 334
 QY 241 AAAGACAGTCAATTTTGAACATCCAGATGCTTTTGAATATGATTTGATGACACAGACT 300
 DB 335 AAAGACAGTCAATTTTGAACATCCAGATGCTTTTGAATATGATTTGATGACACAGACT 394
 QY 301 CTGAAGAACATCGTAGAGGCGCAAAACGCTGAGAGTGCAGCTTATGATTTTGTGACACAC 360
 DB 395 CTGAAGAACATCGTAGAGGCGCAAAACGCTGAGAGTGCAGCTTATGATTTTGTGACACAC 454
 QY 361 TCAAGGTTACAGAGACCAACGCTGTCTACCTCGGACGCTGTCTGTGTTGAGGGCATC 420
 DB 455 TCAAGGTTACAGAGACCAACGCTGTCTACCTCGGACGCTGTCTGTGTTGAGGGCATC 514
 QY 421 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTCCACTGCGGCTCTTGTGTGACACC 480
 DB 515 TTGGTGTCTACAGCGAGAGATCCGGGACATGTTCCACTGCGGCTCTTGTGTGACACC 574
 QY 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCCTCCGGAGCTGCGCCGAGGGAGGACCTTG 540

DB 575 GACTCCGACGTACAGGCTGTCTCGAAGAGTTCCTCCGGAGCTGCGCGAGGAGGACCTTG 634
 QY 541 GAGCAGATTCTGACGAGTACACACACCTTCTGTAAGCCGGCTTGGAGAGTTCGCTTG 600
 DB 635 GAGCAGATTCTGACGAGTACACACACCTTCTGTAAGCCGGCTTGGAGAGTTCGCTTG 694
 QY 601 CCGACAAAGAGTATGCCATGTGATCATCCACGAGAGTGGACATATGTTGTCATC 660
 DB 695 CCGACAAAGAGTATGCCATGTGATCATCCACGAGAGTGGACATATGTTGTCATC 754
 QY 661 AACCTGATGTCGAGACATCCAGACGACATTTGAAATGTGATCATCTGCAATGACACCA 720
 DB 755 AACCTGATGTCGAGACATCCAGACGACATTTGAAATGTGATCATCTGCAATGACACCA 814
 QY 721 GGAGGCTCAATGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGAGACCACTTGGG 780
 DB 815 GGAGGCTCAATGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGAGACCACTTGGG 874
 QY 781 ATGCTGACCTTCGGGCAAGGTCACATTTGAGTCCAGACGAGACCCGACTGA 834
 DB 875 ATGCTGACCTTCGGGCAAGGTCACATTTGAGTCCAGACGAGACCCGACTGA 928
 RESULT 2
 AAD27186
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 AC AAD27186;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 KW Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticonvulsant; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 94..927
 FT /*tag= a
 FT /product= "Human 57658 protein"
 FT misc_feature 94..924
 FT /*tag= b
 FT /note= "This region is specifically claimed as
 FT SEQ ID NO: 3 in claim 1 of the specification"
 XX
 PD WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 DR WPI; 2002-140091/18.
 DR P-PSDB; AAB16592.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX

PS Claim 1; Fig 1a; 103pp; English.

CC The patent discloses human uridine kinase-like polypeptides, designated
CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
CC are useful for developing diagnostic and therapeutic agents for 57658-
CC mediated or related disorders such as haematopoietic neoplastic disorders
CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
CC their antibodies are useful in screening assays, detection assays (e.g.
CC forensic biology) and predictive medicine (e.g. diagnostic assays,
CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
CC They are useful as reagents for diagnosing and treating 57658-mediated
CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping
CC to identify an individual from a minute biological sample (tissue typing)
CC and to aid in forensic identification of the biological sample. The
CC present sequence is a DNA encoding human 57658 protein.

Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Query Match	100.0%	Score 834;	DB 246	Length 1624;
Best Local	100.0%	Pred. NO. 8.9e-206;		
Matches 834;	0;	Mismatches	0;	Gaps 0;
Conservative				

QY	1	ATGGCTCGGCGGAGGCGCAACCTGACGAGAGCCCGCGCGGAGGCGACCGTCCGCAC	60
Db	94	ATGCGCTCGGCGGAGGCGAGACTGTCAGAGAGCCCGCGGAGGCGACCGTCCGCAC	153
QY	61	CAGCGCCCTCTCTGATAGGGGTGAGCGCGCACATGCAAGCGGAAAGTGACCTGTGT	120
Db	154	CAGCGCCCTCTCTGATAGGGGTGAGCGCGCACATGCAAGCGGAAAGTGACCTGTGT	213
QY	121	GAGAAATCATGAGAGTGTGTGGGACAGAAACGAGGTGAAACAGCGGACCGGAAAGTGTCT	180
Db	214	GAGAAATCATGAGAGTGTGTGGGACAGAAACGAGGTGAAACAGCGGACCGGAAAGTGTCT	273
QY	181	ATCTTGAGCGCAGGACAGGTTCTTCAAGGTCTCTGACGCGACAGCAAGAGCGCAAGCCTTG	240
Db	274	ATCTTGAGCGCAGGACAGGTTCTTCAAGGTCTCTGACGCGACAGCAAGAGCGCGCTTG	333
QY	241	AAAGCAGATCAATTTTGAACCATCCAGATGCTTTGATATATGATTTGATGCAAGACT	300
Db	334	AAAGCAGATCAATTTTGAACCATCCAGATGCTTTGATATATGATTTGATGCAAGACT	393
QY	301	CTGAGAAATCATGTGAGGGCAAAACGGTGAAGGTGCGCACTATGATTTTGTGACAC	360
Db	394	CTGAGAAATCATGTGAGGGCAAAACGGTGAAGGTGCGCACTATGATTTTGTGACAC	453
QY	361	TCAGGTTACAGAGACCAACGGTGTCTTACCTTGCGACGTTGTTCTGTTTGAAGGCATC	420
Db	454	TCAGGTTACAGAGACCAACGGTGTCTTACCTTGCGACGTTGTTCTGTTTGAAGGCATC	513
QY	421	TTGGTGTTCACGCGCAGAGAGATCCGGGACATGTTCCACTCGCGCCTCTTCGTGACAC	480
Db	514	TTGGTGTTCACGCGCAGAGAGATCCGGGACATGTTCCACTCGCGCCTCTTCGTGACAC	573
QY	481	GACTCCGACGTCAAGCTGTCTTGAAGAATTCTCCGGACGTCGCGCGAGGAGAGGACCTG	540
Db	574	GACTCCGACGTCAAGCTGTCTTGAAGAATTCTCCGGACGTCGCGCGAGGAGAGGACCTG	633
QY	541	GAGCAGATTCAGACGACATGACCACTTCCTGTAAGCGCGCTTCGAGAGATTCTGCTG	600
Db	634	GAGCAGATTCAGACGACATGACCACTTCCTGTAAGCGCGCTTCGAGAGATTCTGCTG	693
QY	601	CCGACAAAGATATGCGCATGTGATCAATCCACGAGAGGTGACAATATGTTGTCATC	660
Db	694	CCGACAAAGATATGCGCATGTGATCAATCCACGAGAGGTGACAATATGTTGTCATC	753
QY	661	AACCTGATGTGACAGCAATCCAGACATTCTGAATGTGACATCTGCAAATGACACCA	720
Db	754	AACCTGATGTGACAGCAATCCAGACATTCTGAATGTGACATCTGCAAATGACACCA	813

QY	721	GGAGGCTCCATTGGCGGAGCTACAAGGGGACCTTTTCGACGACGAGGGACCAACCTTGGG	780
Db	814	GGAGGCTCCATTGGCGGAGCTACAAGGGGACCTTTTCGACGACGAGGGACCAACCTTGGG	873
QY	781	ATGCTGACCTCTGGCAAAACGGTCACATTTTGGAGTCCAGACGACGACCCCACTGGA	834
Db	874	ATGCTGACCTCTGGCAAAACGGTCACATTTTGGAGTCCAGACGACGACCCCACTGGA	927

RESULT 3
AAH15621

AC AAH15621;

DT 26-JUN-2001 (first entry)

DE Human CDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens

PN EP1074617-A2

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126
VY

PR	29-JUL-1999;	99JP-0248036
PR	27-AUG-1999;	99JP-0300353

PR	11-JAN-2000; 2000JP-0118776
PP	03-MAY-2000; 2000JP-0183767

PR 09-JUN-2000; 2000JP-0241899
VY

PA (HELI-) HELIX RES INST.
VY

PI Ota T, Isogai T, Nishikawa
PI Ichii S, Sugiyama T, Wakai

XX
WDT : 2001-310740/24

XX
DE

full-length cDNAs de-

PT	full-length cDNAs -
xy	

PS Claim 8; SEQ ID 1395

CC The present invention

CC comprises: (a) an ol

CC the 5602 nucleotide

of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AA929446 to CC AA958993 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 2160 BP: 457 A; 591 C; 671 G; 441 T; 0 other;

Query Match 100.0%; Score 834; DB 22; Length 2160;

Best Local Similarity 100.0%; Pred. No. 9.8e-206;

Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCGCGAC 60
DB 95 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCGCGAC 154
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCGCACTGCGAGCGGAAAGTCAGCCGTGTGT 120
DB 155 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCGCACTGCGAGCGGAAAGTCAGCCGTGTGT 214
QY 121 GAGAGATCATGAGTGTCTGGGACAGAACAGAGTGGAAACGCGGACGCGGAAGTGTGTC 180
DB 215 GAGAGATCATGAGTGTCTGGGACAGAACAGAGTGGAAACGCGGACGCGGAAGTGTGTC 274
QY 181 ATCTGAGCCGAGCAGGTTCTACAGGTTCTGACGCGAGAGCAGAAAGGCCAGGCTTG 240
DB 275 ATCTGAGCCGAGCAGGTTCTACAGGTTCTGACGCGAGAGCAGAAAGGCCAGGCTTG 334
QY 241 AAAGCAGATCAATTTTGAACATCCAGATGCTTTGATTAATGATTTGATGACAGGACT 300
DB 335 AAAGCAGATCAATTTTGAACATCCAGATGCTTTGATTAATGATTTGATGACAGGACT 394
QY 301 CTGAAGAACATGTTGAGAGGCGAAGACGCTGAGAGTCCGACCTTATGATTTGTGACAC 360
DB 395 CTGAAGAACATGTTGAGAGGCGAAGACGCTGAGAGTCCGACCTTATGATTTGTGACAC 454
QY 361 TCAAGTTACAGAGACCGAGGTGTCTACCTGCGAGCGGTGCTGTTGAGGGGATC 420
DB 455 TCAAGTTACAGAGACCGAGGTGTCTACCTGCGAGCGGTGCTGTTGAGGGGATC 514
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCTTCTGTGTGACACC 480
DB 515 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCTTCTGTGTGACACC 574
QY 481 GACTTCGACGTCAGGCTGTCTCGAAGAGTTTCCGGGACGTGCGCCGAGGAGGAGGACTTG 540
DB 575 GACTTCGACGTCAGGCTGTCTCGAAGAGTTTCCGGGACGTGCGCCGAGGAGGAGGACTTG 634
QY 541 GAGAGATTTGAGCGCAGTACCAACCTTCGTAAGCGCGCTTCAAGAGTTTCTGCTG 600
DB 635 GAGAGATTTGAGCGCAGTACCAACCTTCGTAAGCGCGCTTCAAGAGTTTCTGCTG 694
QY 601 CCGACAAAGAGTATGCCGATGTGATCAATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 695 CCGACAAAGAGTATGCCGATGTGATCAATCCACGAGAGTGAACAATATGTTGCCATC 754
QY 661 AACCTGATCTGCGACACATCCAGACATTTGTAATGTGACATCTGCAATGGCACCGA 720
DB 755 AACCTGATCTGCGACACATCCAGACATTTGTAATGTGACATCTGCAATGGCACCGA 814
QY 721 GGAGGGTTCATGAGCGGAGCTACAAAGCGACCTTTTCTGACCAAGGGGACACCCCTGGG 780
DB 815 GGAGGGTTCATGAGCGGAGCTACAAAGCGACCTTTTCTGACCAAGGGGACACCCCTGGG 874
QY 781 ATGCTGACCTTCGCAAAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 834
DB 875 ATGCTGACCTTCGCAAAACGCTCAATTTGGAGTCCAGACAGACCCCACTGA 928

```

RESULT 4
ID AAK98735
AAK98735 standard; DNA; 834 BP.

XX AAK98735;

XX 02-MAY-2002 (first entry)

XX DNA of a human uridine kinase (UDK).

XX

KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localization study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 1..780
FT CDS /tag= a
FT /partial
FT /note= "No stop codon"

PN M0200172963-A2.

PD 04-OCT-2001.

PF 27-MAR-2001; 2001MO-US09663.

PR 27-MAR-2000; 2000US-0536647.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Ho YS, Johnson RK;

PP WPI; 2001-626259/72.

DR P-PSDB; AAO14412.

XX

PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance

PS Claim 5; Page 22-23; 31pp; English.

XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilizing uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example, detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localization studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This polynucleotide sequence represents the DNA
CC of a human uridine kinase of the invention.

XX Sequence 834 BP: 194 A; 219 C; 265 G; 156 T; 0 other;

Query Match 99.8%; Score 832.4; DB 23; Length 834;

Best Local Similarity 99.9%; Pred. No. 1.8e-205;

Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGCGAC 60
DB 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGCGAC 60
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCAACCGTGTGT 120
DB 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGAAAGTCAACCGTGTGT 120
QY 121 GAGAGATCATGAGTGTCTGGGACAGAACGAGTGGAAACGCGGACGCGGAAGTGTGTC 180
DB 121 GAGAGATCATGAGTGTCTGGGACAGAACGAGTGGAAACGCGGACGCGGAAGTGTGTC 180

```


QY	181	ATCCGACCCAGSAGAGGTTCTACAAAGTCCGACGCGAGAGAGAAAGCCAAAGCCTTG	240
Db	181	ATCTGAGCAGAGAGAGGTTCTACAAAGTCTGACGCGAGAGAGAAAGCCAAAGCCTTG	240
QY	241	AAAGGACAGTACAAATTTTGACCATCCAGATGCGCTTGTATTAATGATTTGATGACAGAACT	300
Db	241	AAAGGACAGTACAAATTTTGACCATCCAGATGCGCTTGTATTAATGATTTGATGACAGAACT	300
QY	301	CTGAAGAACATCGTGGAGGGCAAAAACGCTGGAGGTGCCAGCATATTTTGTGACACAC	360
Db	301	CTGAAGAACATCGTGGAGGGCAAAAACGCTGGAGGTGCCAGCATATTTTGTGACACAC	360
QY	361	TCAAGGTTACCAAGACCAACGGTGGCTACCCCTGGCGAGAGTGTCTGTTTGAAGGCAATC	420
Db	361	TCAAGGTTACCAAGACCAACGGTGGCTACCCCTGGCGAGAGTGTCTGTTTGAAGGCAATC	420
QY	421	TTGCTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGGCGCTCTTCCGACAC	480
Db	421	TTGCTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGGCGCTCTTCCGACAC	480
QY	481	GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCCTCCGGGACGTGCGCCGAGGAGGACCTG	540
Db	481	GACTCCGACGTCAAGGCTGTCTCGAAGAGTTCCTCCGGGACGTGCGCCGAGGAGGACCTG	540
QY	541	GAGCAGATTTCTGACGCGAGTACACACCTTCGGTGAACCGCGCCCTTGAGAGATTCGCTG	600
Db	541	GAGCAGATTTCTGACGCGAGTACACACCTTCGGTGAACCGCGCCCTTGAGAGATTCGCTG	600
QY	601	CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCATC	660
Db	601	CCGACAAAGAAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCATC	660
QY	661	AACTGATGCTGACAGACATCCAGGACATTTCTGAATGTGACATCTGCAAAATGGACCGA	720
Db	661	AACTGATGCTGACAGACATCCAGGACATTTCTGAATGTGACATCTGCAAAATGGACCGA	720
QY	721	GGAGGGTCCAAATGGGCGGAGCTACAAAGCGGACCTTTCTGAGCCAGGGGACCAACCTGGG	780
Db	721	GGAGGGTCCAAATGGGCGGAGCTACAAAGCGGACCTTTCTGAGCCAGGGGACCAACCTGGG	780
QY	781	ATGCTGACCTCTGGCAAAAGCTACATTTTGGATGTCAGACGACGACCCCACTGA	834
Db	781	ATGCTGACCTCTGGCAAAAGCTACATTTTGGATGTCAGACGACGACCCCACTGA	834
RESULT 5			
AAI60444			
ID	AAI60444	standard; cDNA; 2152 BP.	
XX	AAI60444;		
XX	AC		
XX	DT	22-OCT-2001 (first entry)	
XX	DE	Human polynucleotide SEQ ID NO 4433.	
XX			
KW	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;		
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;		
XX	leukemia; ss.		
OS	Homo sapiens.		
XX	PN	WO200153312-A1.	
XX	PD	26-JUL-2001.	
XX	PF	26-DEC-2000; 2000WO-US34263.	
XX	PR	21-JAN-2000; 2000US-0488725.	
XX	PR	25-APR-2000; 2000US-0552317.	

PR	09-JUL-2000;	2000US-0598042.	
PR	19-JUL-2000;	2000US-0620312.	
PR	03-AUG-2000;	2000US-0653450.	
PR	14-SEP-2000;	2000US-0662191.	
PR	19-OCT-2000;	2000US-0693036.	
PR	29-NOV-2000;	2000US-0727344.	
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YF, Liu C, Aouni V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QH, Zhou P, Goodrich R, Drmanac RT;		
XX			
DR	WPI: 2001-442253/47.		
XX	P-PSDB; AAM41288.		
PT	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX			
PS	Claim 1; SEQ ID NO 4433; 10078pp; English.		
XX			
CC	The invention relates to human nucleic acids (AA157798-AA161369) and		
CC	the encoded polypeptides (AAM8642-AAM42213) with nootropic.		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilitation of the activities such as: Immune system suppression,		
CC	Activation/inhibit activity, chemotactic/chemokinetic activity, haemostatic		
CC	and chromolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification.		
XX			
XX			
SEQ	Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;		
Query Match	99.8%; Score 832.4; DB 22; Length 2152;		
Best Local Similarity	99.9%; Pred. No. 2.5e-205;		
Matches	833; Conservative 0; Mismatches 1; Indels 0; Gaps 0.		
OY	1 ATGGCTTCGGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGAGCCGATCCGAC	60	
DB	59 ATGGCTTCGGGGAGGCGAAGACTGCGAGAGCCCGCGCGAGAGCCGATCCGAC	118	
OY	61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCGACTGCGAGCGGAACTGACCGTGT	120	
DB	119 CAGCGGCCCTTCTGATAGGGGTGAGCGCGCGACTGCGAGCGGAACTGACCGTGT	178	
OY	121 GAGGAAGATCATGAGATTGCTGGGACAGAACAGAGTGGAAACGCCGACGGAAAGTGGTC	180	
DB	179 GAGGAAGATCATGAGATTGCTGGGACAGAACAGAGTGGAAACGCCGAAAGTGGTC	238	
OY	181 ATCTCGAGCCGAGGACAGGTTTCAACAGGTCCTGACCGGACAGGAGGCAAGGCCCTTG	240	
DB	239 ATCTCGAGCCGAGGACAGGTTTCAACAGGTCCTGACCGGACAGGAGGCAAGGCCCTTG	298	
OY	241 AAGGACAGTATCAATTTTGAACATCCATCCAGATCCCTTGTATATGATTGATGACAGGACT	300	
DB	299 AAGGACAGTATCAATTTTGAACATCCATCCAGATCCCTTGTATATGATTGATGACAGGACT	358	
OY	301 CTGAAGAACATCTGAGAGGCGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACAC	360	
DB	359 CTGAAGAACATCTGAGAGGCGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACACAC	418	
OY	361 TCAAGGTTTACAGAGACCAAGGTGTCTAACCTTCGAGCGAGTGTCTGTTTGAAGGCGATC	420	
DB	419 TCAAGGTTTACAGAGACCAAGGTGTCTAACCTTCGAGCGAGTGTCTGTTTGAAGGCGATC	478	
OY	421 TTGGTGTCTTACACCCAGAGATCCGGGACATGTTTCCACTCGCCCTTGTGTGGACAC	480	

```

Db      479 TTGGTGTCTACAGCCAGAGATCGGAGACATGTTCCACTGCGCTTCTTGAGACACC
Qy      481 GACTCCGACGTGAGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG
Db      539 GACTCCGACGTGAGCTGTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG
Qy      541 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAGAGTTCTGCCTG
Db      599 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAGAGTTCTGCCTG
Qy      601 CCGACAAAGAGTATGCGGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATC
Db      659 CCGACAAAGAGTATGCGGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATC
Qy      661 AACCTGATCTGACGACATCGAGACATTTCTGATGATGATCTGCAATTTGGACCGA
Db      719 AACCTGATCTGACGACATCGAGACATTTCTGATGATGATCTGCAATTTGGACCGA
Qy      721 GGAAGGTTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG
Db      779 GGAAGGTTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG
Qy      781 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA
Db      839 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA

```

RESULT 6

ABL89762
ID ABL89762 standard; cDNA; 1396 BP.

AC ABL89762;

DT 24-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 324.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
XX antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX PD 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX PR 19-MAY-2000; 2000US-205515P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI: 2002-122018/16.

XX DR P-PSDB: ABB89353.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders -

XX PS Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (ABB89040-ABB90444) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful

CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;

Query Match 97.0%; Score 808.8; DB 24; Length 1396;
Best Local Similarity 99.4%; Pred. No. 2.8e-199;
Matches 829; Conservative 3; Mismatches 0; Indels 2; Gaps 2;

```

Qy      1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCCGACCTCCGAC
Db      41 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGGCCGACCTCCGAC
Qy      61 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAATGACCTGTGT
Db      101 CAGCGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAATGACCTGTGT
Qy      121 GAGAAGTCAATGAGTTCCTGGGACAGAACAGAGTGGAAAGGCGGAGAGTGTGTC
Db      161 GAGAAGTCAATGAGTTCCTGGGACAGAACAGAGTGGAAAGGCGGAGAGTGTGTC
Qy      181 ATCCTGAGCCAGAGACAGTTCCTCAAGGTCCTACGCGCAGAGCAGAAAGGCCAGCTTG
Db      221 ATCCTGAGCCAGAGACAGTTCCTCAAGGTCCTACGCGCAGAGCAGAAAGGCCAGCTTG
Qy      241 AAGGACAGTACAAATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT
Db      281 AAGGACAGTACAAATTTGACATCCAGATGCTTTGATATGATTTGATGACAGAGCT
Qy      301 CTGAAGAACATCGTGGAGGGCAAAACGTCGAGAGTCCGACCTATATTTTGTGACAC
Db      341 CTGAAGAACATCGTGGAGGGCAAAACGTCGAGAGTCCGACCTATATTTTGTGACAC
Qy      361 TCAAGTTACAGAGACCAAGTGTCTACCTCGCGAGCGTGTCTGTTTGGAGGCATC
Db      401 TCAAGTTACAGAGACCAAGTGTCTACCTCGCGAGCGTGTCTGTTTGGAGGCATC
Qy      421 TTGGTGTCTACAGCAAGAGATCCGAGACATGTCACCTGCGCTCTTCTGTGACACC
Db      461 TTGGTGTCTACAGCAAGAGATCCGAGACATGTCACCTGCGCTCTTCTGTGACACC
Qy      481 GACTCCGAGCGTCAAGCGTGTCTGAAAGATTCTCCGAGAGTGGACGAGGAGGACCTG
Db      521 GACTCCGAGCGTCAAGCGTGTCTGAAAGATTCTCCGAGAGTGGACGAGGAGGACCTG
Qy      541 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAGAGTTCTGCCTG
Db      581 GAGCAGATTCTGACGACGATACCACTTCTGTAAGCGGCGCTTGAGAGTTCTGCCTG
Qy      601 CCGACAAAGAGTATGCGGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATC
Db      640 CCGACAAAGAGTATGCGGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATC
Qy      661 AACCTGATCTGACGACATCGAGACATTTCTGATGATGATCTGCAATTTGGACCGA
Db      700 AACCTGATCTGACGACATCGAGACATTTCTGATGATGATCTGCAATTTGGACCGA
Qy      721 GGAAGGTTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG
Db      760 GGAAGGTTCAATGGCGGAGCTACAGCGGACCTTTTCTGAGCGCAAGGGACCACTGGG
Qy      781 ATGCTGACCTCTGGCAACCGTCACTTTGAGAGTCCAGACAGACCCCACTGA

```

Db 819 ATGCTGACCTCTGCAAAACGTCACATTTGGAGTCCAGACGACAGACCCCACTGA 872

|||||

RESULT 7
AAH75355
AAH75355 standard; cDNA; 900 BP.

XX AAH75355;
XX
XX 02-OCT-2001 (first entry)
XX
XX Human uridine kinase encoding cDNA.
XX
XX Human, uridine kinase; UK; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 8..838
XX FT /*tag= a
XX FT /product= "uridine kinase"
XX FT /note= "claimed in claim 1"
XX
XX CN1287172-A.
XX
XX 14-MAR-2001.
XX
XX 07-SEP-1999; 99CN-0118818.
XX
XX 07-SEP-1999; 99CN-0118818.
XX
XX (UYFU-) UNIV FUDAN.
XX
XX Yu L, Zhao Y, Zhang H;
XX
XX WPI; 2001-409529/44.
XX P-PSDB; AAG64506.
XX
XX Human uridine kinase and its coding sequence, preparation and
XX application -
XX
XX Claim 1; Page 15 (Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;
XX
XX Query Match 94.2%; Score 786; DB 22; Length 900;
XX Best Local Similarity 97.2%; Pred. No. 1.9e-193;
XX Matches 811; Conservative 0; Mismatches 20; Indels 3; Gaps 1;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGACCCCGCGGAGCGCGACCGTCCGAC 60
DB 8 ATGGCTTCGGCGGAGGCGAAG--GTGCGAGAGCCGCGCGCGAGCGAACCGTCCGAC 64

QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCGTGT 120
DB 65 CAGCGTCCCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCGTGT 124

QY 121 GAGAGATCATGAGTGTCTGGGACGAGAGAGGTGGAAGAGCGGCGCGGAAGTGTCTC 180
DB 125 GAGAGATCATGAGTGTCTGGGACGAGAGAGGTGGAAGAGCGGCGCGTAAGTGTCTC 184

QY 181 ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGCAGAGGAGAGGCGGCTTG 240
DB 185 ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGCAGAGGAGGCGGCGCTTG 244

QY 241 AAGAGACAGTACATTTTGAACATCCAGATGCGCTTTGATATGATTTGATGACAGGACT 300
DB 245 AAGAGACAGTACATTTTGAACATCCAGATGCGCTTTGATATGATTTGATGACAGGACT 304

QY 301 CTGAGAACATCTGTGAGGCGCAAAACGCGTGAAGGTGCCGACTATGATTTTGTGACACAC 360

Db 305 CTGAGAACATCTGTGAGGCGCAAAACGCTGAGAGGTGCCGACCTATGATTTTGTGACACAC 364

QY 361 TCAAGGTACACAGACACACGCTGTCTACCCCTGCGGAGGTGTTCTGTTTGAAGGCATC 420
DB 365 TCAAGGTACACAGACACACGCTGTCTACCCCTGCGGAGGTGTTCTGTTTGAAGGCATC 424

QY 421 TTGTTGTTCTACAGCAGGAGGATCCGGGACATGTTCCACTGCGCTCTTGTGAGACAC 480
DB 425 TTGTTGTTCTACAGCAGGAGGATCCGGGACATGTTCCACTGCGCTCTTGTGAGACAC 484

QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTTCTCCGGGACGTGCGCCGAGGAGGACCTG 540
DB 485 GACTCCGAGCTCAGGCTGTCTCGAAGATTTCTCCGGGACGTGCGCCGAGGAGGACCTG 544

QY 541 GAGCAGATTTCTGACGACATCACACCTTGTGAGACCGGCTTTCGAGAGATTCTGCTG 600
DB 545 GAGCAGATTTCTGACGACATCACACCTTGTGAGACCGGCTTTCGAGAGATTCTGCTG 604

QY 601 CCGACAAAGAGATGCCGATGATCATCCGACGAGAGTGGACAAATAGTTGTCATC 660
DB 605 CCGACAAAGAGATGCCGATGATCATCCGACGAGAGTGGACAAATAGTTGTCATC 664

QY 661 AACCTGATGTCGACGACATCCAGACATTTGATGATGTCATGCAATGSGCACCGA 720
DB 665 AACCTGATGTCGACGACATCCAGACATTTGATGATGTCATGCAATGSGCACCGA 724

QY 721 GAGGAGTCCATAGTGGCGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGACACCTGG 780
DB 725 GAGGAGTCCATAGTGGCGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGACACCTGG 784

QY 781 ATGCTGACCTCTGCGCAACGCTCACATTTGGAGTCCAGACGAGACCCCACTGA 834
DB 785 ATGCTGACCTCTGCGCAACGCTCACATTTGGAGTCCAGACGAGACCCCACTGA 838

RESULT 8
AAH04832
AAH04832 standard; cDNA; 753 BP.

XX AAH04832;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:1667.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EPI074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000BP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -

Best Local Similarity 88.6%; Pred. No. 3.1e-115;
Matches 558; Conservative 0; Mismatches 10; Indels 62; Gaps 1;

```

QY 260 ACCATCCAGATGCTTGTATATGATTTGATGACAGAGACTCTGAAAGACATGTGGAGG 319
DB 6 ACCCTGAGATGCTTGTATATGATTTGATGACAGAGACTCTGAAAGACATGTGGAGG 65
QY 320 GCAAAAACGATGAGAGTCCGACCTATGATTTTGTGACACCTAAGGTTCCAGAGACA 379
DB 66 GCAAAAACGATGAGAGTCCGACCTATGATTTTGTGACACCTAAGGTTCCAGAGACA 125
QY 380 CGGTGTCTACCTCGGAGAGTGTCTGTGAGGACATCTTGTTCTACAGCCAGG 439
DB 126 CGGTGTCTACCTCGGAGAGTGTCTGTGAGGACATCTTGTTCTACAGCCAGG 185
QY 440 AGATCCGGGACATGTTCCATCTCGCTCTTCTGTGACACCGACTCCGAGCTGAGCTGT 499
DB 186 AGATCCGGGACATGTTCCATCTCGCTCTTCTGTGACACCGACTCCGAGCTGAGCTGT 245
QY 500 CTGGAAGAGTCTCCGGGAGAGTGTGCGCGAGAGAGGACCTGAGACATCTGACCGAGT 559
DB 246 CTGGAAGAGTCTCCGGGAGAGTGTGCGCGAGAGAGGACCTGAGACATCTGACCGAGT 305
QY 560 ACAACCACTTCTGTGAAAGCCGCTTCCAGGAGTCTTCTGCTGCGG----- 603
DB 306 ACAACCACTTCTGTGAAAGCCGCTTCCAGGAGTCTTCTGCTGCGGACAGACATCTGAC 365
QY 604 -----ACAAAGAGTATGC 617
DB 366 AGGGAATGAGAGTACAGCATTTGAGCAATGAGTGTGATGAGGGAACAAAGATATGC 425
QY 618 CGATGTGATCATCCACGAGAGTGAACAATATGTTGATCACTGATCTGTGACGA 677
DB 426 CGATGTGATCATCCACGAGAGTGAACAATATGTTGATCACTGATCTGTGACGA 485
QY 678 CATCCAGACATTTCTGATGATGATGATCTGCAATGAGACCGAGAGAGTCTCAATGGCG 737
DB 486 CATCCAGACATTTCTGATGATGATGATCTGCAATGAGACCGAGAGAGTCTCAATGGCG 545
QY 738 GAGCTACAGCGGACCTTTCTGAGACCGAGAGACACCTGAGAGTGTGATCTGTGACAA 797
DB 546 GAGCTACAGCGGACCTTTCTGAGACCGAGAGACACCTGAGAGTGTGATCTGTGACAA 605
QY 798 ACGGTACATTTGAGAGTCCAGAGACAC 827
DB 606 ACGGTACATTTGAGAGTCCAGAGACACCTGAGAGTGTGATCTGTGACAA 635

RESULT 10
AAH23801
ID AAH23801 standard; cDNA; 1322 BP.
XX
XX AAH23801;
AC
AC AAH23801;
DT
DT 31-JUL-2001 (first entry)
XX
XX Human transferase HTFS-1 cDNA, SEQ ID NO:43.
DE
DE Human transferase; HTFS; agonist; antagonist; cellular signalling;
XX Human proliferation; cell proliferative disorder; immune disorder;
XX atherosclerosis; hepatitis; psoriasis; cancer; tumour;
XX inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
XX cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
XX multiple sclerosis; rheumatoid arthritis; pancreatitis;
XX systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
XX haemodialysis; extracorporeal circulation; trauma; transgenic animal;
XX gene therapy; drug screening; ss.
XX
XX Homo sapiens.
OS
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT CDS 250..1025
FT /tag= a

```

```

PT /product= "HTFS-1"
FT /function= "Transferase"
XX
XX WO200132888-A2.
XX
XX 10-MAY-2001.
XX
XX 02-NOV-2000; 2000MO-US30485.
XX
XX 04-NOV-1999; 99US-0163595.
XX
XX (INCYTE GENOMICS INC.
XX
XX Tang YF, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
XX Shi L, Azimzai Y, Lu DM, Baugh MR;
XX WPI; 2001-328796/34.
XX P-PSDB; AAB73494.
XX
XX Human transferase polypeptides and polynucleotides useful for
XX diagnosis, prevention and treatment of cell proliferative and immune
XX system disorders and for identifying agonists and antagonists -
XX
XX Claim 5; Page 135-136; 157pp; English.
XX
XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
XX HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
XX encoding them. The proteins play important roles in the regulation of
XX cellular signalling and proliferation. The HTFS proteins are useful for
XX screening compounds for their effectiveness as agonists or antagonists of
XX transferase activity, or for compounds that specifically bind to an HTFS
XX protein or which modulates the activity of an HTFS protein.
XX
XX Pharmaceutical compositions comprising an HTFS protein, HTFS
XX agonist or antagonist, or genetic construct encoding an HTFS
XX protein are useful for treating a disease or condition associated
XX with decreased or increased expression of functional HTFS. Disorders
XX which may be treated using such compositions include cell proliferative
XX disorders and immune disorders. For example, diseases which may be
XX treated include atherosclerosis, hepatitis, psoriasis, cancers (including
XX breast, bladder, bone marrow, brain and uterus cancer), inflammation,
XX AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
XX disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
XX rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
XX thrombocytopenia, and ulcerative colitis. They may also be used to treat
XX complications of cancer, haemodialysis, extracorporeal circulation,
XX trauma and haematopoietic cancer, including lymphoma, leukaemia and
XX myeloma. Polynucleotides encoding HTFS proteins are useful for creating
XX transgenic animals to model human diseases, for diagnostic purposes and
XX to generate hybridisation probes useful in mapping the naturally
XX occurring genomic sequences. HTFS, and its catalytic or immunogenic
XX fragments are useful for screening libraries of compounds in a variety of
XX drug screening techniques. Antibodies which specifically bind HTFS may be
XX used for the diagnosis of disorders associated with the expression of
XX HTFS, or in assays to monitor patients being treated with HTFS or
XX agonists, antagonists or inhibitors of HTFS. The present sequence
XX represents an HTFS protein-encoding cDNA of the invention.
XX
XX Sequence 1322 BP; 339 A; 359 C; 358 G; 266 T; 0 other;
XX
XX Query Match 44.8%; Score 374; DB 22; Length 1322;
XX Best Local Similarity 74.3%; Pred. No. 8.6e-87;
XX Matches 486; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 66 GCCCTTCTGATAGGGGTGAGCGCGGACCTCCAGCGGGAAGTCAGCCGTGTGAGAA 125
DB 306 GCCCTTCTTATATAGGGGTGAGCGGGGAACTAGCGGCAAGTCTTCGTGTGTCTAA 365
QY 126 GATCATGAGTGTCTGTGAGACAGAGAGTGAACAGCGGACGGAAGGTGTCTCT 185
DB 366 GATCTGTGAGCTCTGTGGGCAAGATGAGGTGATCTATGCGCAAGACAGGTGTCTCT 425
QY 186 GAGCCAGAGACAGTCTCAAGAGTCTGTGACGCGAGAGCCAGAGGCTTGAAGG 245

```

Db	426	GAGCCAGGATACCTTCTACCGGTCTCTTACCTCGGAGCAGAAGGCCAAGACCTCTGAAGG	485
Oy	246	ACAGTACAAATTTTGACCATTCAGATCCCTTTGATAATGATTGATGCA CAGACTGTAA	305
Db	486	CCAGTTCAACTTTTGACCAACCGGATGCTTTGCAATGAACCTATTCCTCAAAACACTCA	545
Oy	306	GAACATCGTGAGGGGCAAAACGGTGGAGGTCGCAGCTTGATTTTGTGCA CACTCAAG	365
Db	546	AGAAATCACTGAAGGAAAAACAETCCAGATCCCCGTGTATGACTTTGTCTCCATTCGG	605
Oy	366	GTTACCAAGAGACCAACGGTGTCTACCCCTGCGACGTGTCTCTGTTTGAAGGCACTTGGT	425
Db	606	GAAGGAGGAGACAGTTACTGTCTATCCCGACAGCTGTGTCTTTGAAGGAGCTCTGGC	665
Oy	426	GTTCTACAGCCGAGGATCCGGGACAAATGTTCCACCTGGCCCTTGTGTGACACCGACATC	485
Db	666	CTTCTACTCTCCACGAGAGTACGAGACTGTGTTCCAGATGAACCTTTTGTGTGATACAAATGC	725
Oy	486	CGACGTCAAGCTGTCTCGAAGAGTTCTCCGGACGTGCGC -- CGAAGGAGGACCTTGA	542
Db	726	GGACACCGGGCTCTCCGACAGATTAAGAAGGCATCAAGGAGAGAGGCGGGAATCTTGA	785
Oy	543	GCAGATTTCTGAGCGAGTATACCACTTGTGTGAAGCCGGCTTGSAGAGTTCTGTGCTGGC	602
Db	786	GCAGATTTTATCTCGATACATTACGTTGCTGTAAGCCCTGCTTTGAGAGAAATTCGTGTGGC	845
Oy	603	GACAAAGAAGTATGCGATGTATCATCCACAGAGAGTGAACAATATGTGTTCCATCAA	662
Db	846	AACAAAGAAGTATGCTGATGTATATATCCCTGAAGTGCAGAAATCTGTGTGCATCAA	905
Oy	663	CTGATCTGTGACGACATCCAGACATTTTGATGTGTACATCTGCAAAATGCA	716
Db	906	CTCATCTGTGACGACATCCAGAGACATCTGTGATGAGGGGCCCTTCAAAACGGCA	959

RESULT 11	
AA157850	
ID	AA157850 standard; cDNA, 1402 BP.
XX	
AC	AA157850;
XX	
DT	22-OCT-2001 (first entry)
DE	Human polynucleotide SEQ ID NO 53.
XX	
KW	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW	peripheral nervous system; neuropathy; central nervous system; CNS;
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;
KM	leukemia; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200153312-A1.
XX	
PD	26-JUL-2001.
XX	
PF	26-DEC-2000; 2000MO-US34263.
XX	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0598042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
P1	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
I1	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao Qx, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPT: 2001-442253/47.
DR P-PSDB; AAM38694.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Claim 1; SEQ ID NO 53; 10078bp; English.
PS

CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA358642-AA442213) with neurotropic, and
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemoclastic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

Sequence 1402 BP; 379 A; 370 C; 371 G; 282 T; 0 other;

Query Match	44.8%	Score 374;	DB 22;	Length 1402;
-------------	-------	------------	--------	--------------

Matches 486; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

OY	66	GCCCTTCTGATAGGGGGAGCGGGCGGACATGCGCCGAGGGAAGTGAACCGTGTGGAGA	125
Db	346	GCCCTTCTTATAGGCGCTCAGCGGGGGAGCAAGCTAGCGGCAAGTCTCCGTGTGTGCTRA	405
OY	126	GATCATGAGTTGCTGGACAGAAAGAGTGGAAACAGCGGACGCGGAAGTGGTATCCT	185
Db	406	GATGCTGAGCTCCTGGGGCAGAAATGAGGTGGAATATGGCCAGAAAGAGGTGATCTCT	465
OY	186	GAGCCAGGACAGGTTCTACAAAGTCTGACGGCAGAGCAAAAGGCCAAGCCTTGAAAG	245
Db	466	GAGCCAGGATGCTTCTACCGTGTCTTAACTCCGAGAGCAAGAGCCAAAGCCCTGAAGG	525
OY	246	AAGTACAAATTTTGAACAATCAGATAGCCTTTGATATATGATTTGATGACAAGSACTGTAA	305
Db	526	CAAGTTCACTTTGACCAACCGGAGTCCCTTTGACATATACATCAATCTTCAAAAACCTGAA	585
OY	306	GAACATCGTGAGGGCAAAACGGTGAAGTCCCGCATGATATTTGTGACACATCTGAAG	365
Db	586	AGAAATCACTGAAGGAAAAAAGTCCCAATCCCGTGTATATGACTTTGTCTCCATTCGG	645
OY	366	GTTACCAAGAACCAACGTTGTCTACCTTCCGAGCGTGTCTTGTTGAGGCAATCTTGGT	425
Db	646	GAAGGAGAGACAGTTACTGTCTATCCCGCAGACGTTGTCTCTTTGAAGGGAATCTTGGC	705
OY	426	GTTCTACAGCCAGGAGATCCGGGACATGTTCAACTGGCCTCTTCCGAGACACCGACTC	485
Db	706	CTTCTACTCCAGAGAGGTACGAGACCTGTTCCAGATGAAGCTTTTGTGGATACAGATGC	765
OY	486	CGACGTCAGCGCTGTCTCGAAGAGTTCTCCGGGACGTGGC--CGAGGGAAGGACCTTGA	542
Db	766	GGACACCCGGCTCTCACGCAAGATTTAAAGGACATACGCAAGAGGCAAGGATCTTGA	825
OY	543	GCAGATTTCTAGCAGATACACACACTTTCGTGAAGCCGGCCTTCCAGAGATTCTGCTTGC	602
Db	826	GCAATTTTATCTCAGTACATTAAGTTCGTCAAGCCGTGCTTTGAGGAATCTGCTTGC	885
OY	603	GACAAAGAGATATGCCATGTGATCATCTCCACAGAGATGGACAATATGTTGCCATCAA	662
Db	886	AACAAAGAGATATCTGATGTGATCATCTCCATGAGAGGTGCACATATATCTGGTGGCATCAA	945

QY 663 CCTGATGTCAGACATCCAGACATTTGATGTGACATCTGCAATGCGCA 716
 DB 946 CCTGATGTCAGACATCCAGACATCTGATGATGAGGGCCCTCCAAACGGCA 999

RESULT 12

AAF15785
 ID AAF15785 standard; cDNA, 1310 BP.

XX AAF15785;

XX 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:220.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KM vulnoproductive; cytoskeletal; cardioprotective; immunomodulatory; muscular;
 KM vulnereary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
 KM antibacterial; gene therapy; neural; immune; reproductive; renal;
 KM gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KM wound; infectious disease; ss.

XX Homo sapiens.

XX W0200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX P-PSDB; AAB56582.

PT Prostate cancer associated gene sequences, referred to as prostate
 cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

XX Claim 1; Page 752; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
 CC cardioprotective, immunomodulatory, muscular, vulnereary, gastrointestinal,
 CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotide may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF15566 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present

CC invention.

XX Sequence 1310 BP; 338 A; 350 C; 352 G; 264 T; 6 other;

XX Query Match 43.4%; Score 361.6; DB 21; Length 1310;

XX Best Local Similarity 74.0%; Pred. No. 1,4e-83;

XX Matches 484; Conservative 1; Mismatches 165; Indels 4; Gaps 2;

QY 66 GCCCTTCCTGATGAGGCGGCACTGCAGCGGGAAGTCCAGCGTGTGAGAA 125

DB 288 GCCCTTCCTGATGAGGCGTCA-CGGGGGAACAGCTAGCGGCAAGCTTCCGCTGTGCTAA 346

QY 126 GATCATGAGAGTCTGTGGACAGAACGAGTGAACAGCGGACGGAAGGTGTCACTCT 185

DB 347 GATCTGTGACAGCTCTCGGGGAGATGAGTGTGACTATCCGACAGAGGTGTCACTCT 406
 QY 186 GACCCAGACAGGTTTCTTCAAGAGTCTGTGACGGAGAGACAGAAAGCCCTTTGAAAG 245
 DB 407 GAGCCAGAGTAAAGCTTCTTCAAGAGTCTGTGACGGAGAGACAGAAAGCCCTTTGAAAGS 466
 QY 246 ACAGTACATTTTGAACATCCAGATCCAGATCTTGTATTAATGATTTGATGACAGACTGAA 305
 DB 467 CCACTTCACTTGAACATCCAGATCCAGATCTTGTATTAATGATTTGATGACAGACTGAA 526
 QY 306 GAAATCATCTGAGAGGAGAAACAGTGTGAGAGTCCGACCTATGATTTTGTGACAGACTGAA 365
 DB 527 AGAATCATCTGAGAGGAGAAACAGTGTGAGAGTCCGACCTATGATTTTGTGACAGACTGAA 586
 QY 366 GTTACCAAGACCAAGTGTGTCTACCTTGTGAGAGTCTTGTGAGAGTCTTGTGAGAGTCTTGT 425
 DB 587 GAAAGAGAGACAGTGTGTCTACCTTGTGAGAGTCTTGTGAGAGTCTTGTGAGAGTCTTGTGAG 646
 QY 426 GTTCTACAGCAGAGATCCGAGACATGTTCCAGCTGCGCTCTGTGAGAGCAGACTG 485
 DB 647 CTTCTACTCCAGAGAGTACAGAGCTGTTCAGATGAAGCTTTTGTGATACAGATGC 706
 QY 486 CGACGTCAAGCTGTCTGAAAGTCTCCGAGACGTGCGC---CGAGGAGGAGACTGGA 542
 DB 707 GAGACACCGGCTCTCAAGAGATTTAAGGACATCAGAGAGAGGAGGAGATCTTGA 766
 QY 543 GCGATTTCTGACGAGTACCACTCTTGTGAAGCCGCTTGTGAGAGTCTTGTGAGAGTCTTGTGAG 602
 DB 767 GCGATTTTATCTCAAGATCACTTACCTGTCAGAGCTGCTTGTGAGAGTCTTGTGAGAGTCTTGTGAG 826
 QY 603 GACAAAGAGTATGCGATGTGATCCAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 662
 DB 827 AACAAAGAGTATGCTGATGTGATCTCTGTAAGGTGACAGATATCTGTGAGAGTGTGAGAGTGTGAG 886
 QY 663 CCTGATGTCAGACATCCAGACATTTGATGTGACATCTGCAATGCGCA 716
 DB 887 CCTGATGTCAGACATCCAGACATCTGATGATGAGGGCCCTCCAAACGGCA 940

RESULT 13

ABQ54470

ID ABQ54470 standard; cDNA, 1322 BP.

XX ABQ54470;

XX 22-AUG-2002 (first entry)

DE Human ovarian antigen HOBLP29 cDNA, SEQ ID NO:350.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive; chromosome 1p32;
 KM gene; ss.

XX Homo sapiens.

XX W0200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US16569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;
 XX WPI: 2002-147878/19.
 DR P-PSDB; ABP41393.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 1; SEQ ID NO 350; 2922bp; English.

XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX

SQ Sequence 1322 BP; 351 A; 351 C; 352 G; 264 T; 4 other;

Query Match 43.4%; Score 361.6; DB 24; Length 1322;
 Best Local Similarity 74.0%; Pred. No. 1.4e-83;
 Matches 484; Conservative 1; Mismatches 165; Indels 4; Gaps 2;

QY 66 GCCCTTCTGATAGGGGTGAGCGGCGACCTGCCAGCGGAGATCGACCTGTGAGAA 125
 |||||
 DB 288 GCCCTTCTTATAGGCGTCA-CGGGGGAACAGTACGGCGAGCTTCCGTGTGCTAA 346
 |||||
 QY 126 GATATGAGATGTTGCTGGGACGAACGAGGTGAACAGCGGACGGAGGTGATCTT 185
 |||||
 DB 347 GATGTGACAGTCTCTGGGGGAGATGAGTGAATCTATCGCAGAGAGGTGCTATCT 406
 |||||
 QY 186 GAGCCAGAGACGTTCTTCAACAGTCTGACGGCAGACAGAGGCCAAGGCTTGAAGG 245
 |||||
 DB 407 GAGGACGAGATAGCTTCTTACCTGCTTCACTCGAGAGAGAGGCCAAGGCTTGAAGG 466
 |||||
 QY 246 ACAGTACAAATTTTACCATCCAGATGCTTGTATATGATTTTGAACAGACTTGA 305
 |||||
 DB 467 CCAATTCACATTTTACCAACCCGGATGCTTTGACAAATGAATCTATTTCAAAACACTCAA 526
 |||||
 QY 306 GACATGCTGAGGCGCAAAACGATGAGTGGACCGACCTATGATTTTGAACACTCAAG 365
 |||||
 DB 527 AGAATACATGTAAGGAGAAACAGTCCAGATCCCGTATGATCTTGTCCCATTCGCG 586
 |||||
 QY 366 GTTACGAGAGCCACGATGCTTACCTCGGAGAGTGTGTTTGAAGGCACTTGGT 425
 |||||
 DB 587 GAAGGAGAGACAGTACTGCTATCTTATCCGAGAGAGTGTGCTTTGAAGGAGCTCTGGC 646

QY 426 GTTCTACAGCCAGAGATCCGGACAGTGTCCACCTGCGCTTCTGAGACCCGACTC 485
 |||||
 DB 647 CTCTTACTCCACAGAGGTACGAGACCTGTTCCAGATGAACTTTTGTGATACAGATGC 706
 |||||
 QY 486 CGACGTGAGGCTGTCTTGAAGAGTTCTCCGGACGTGCGC--CGAGGAGGAGACTTGA 542
 |||||
 DB 707 GGAACCCGCGCTCTCAGCGCAGAGTATTAAGGAGCATCAGCAGAGAGGCGAGGATCTTGA 766
 |||||
 QY 543 GCATTTCTGACCGCAGTACCAACCTTCGTGAAGCCGCGCTTGAAGAGTCTTCCCTGCC 602
 |||||
 DB 767 GCAGATTTTATTCAGTACATTACGTTCTGCAAGCCCTGTTGAGGAAATCTGCTTGGC 826
 |||||
 QY 603 GACAAAGAAATATCCGATGTGATCATCCACGAGAGTGAACATATGTTGCCATCAA 662
 |||||
 DB 827 AACAAAGAAATATGCTGATGTGATCATCCCTAAGGTGCGAATATCTGTTGSCCATCAA 886
 |||||
 QY 663 CCGTATGCTGACACATCCAGACATTTCTGAATGTGACATTTGCAATGCGA 716
 |||||
 DB 887 CCGTATGCTGACACATCCAGACATCTGAAATGAGGCGCTCCAAACGCGCA 940
 |||||

RESULT 14

ABL10981
 ID ABL10981 standard; cDNA; 822 BP.
 XX
 AC ABL10981;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 27425.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 KW
 XX Drosophila melanogaster.
 OS
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI: 2001-656860/75.
 DR P-PSDB; ABB66878.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 PS Claim 1; SEQ ID NO 27425; 21bp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU1840-ABU16175) and the encoded proteins
 CC (ABBS7737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 822 BP; 219 A; 219 C; 217 G; 167 T; 0 other;

Query Match 36.2%; Score 302; DB 23; Length 822;

This Page Blank (uspto)


```

Db      181 ATCTGAGCCAGACAGTTCACAGTCTCTGACGCGACAGACGAAAGGCCAGCTTG 240
QY      241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
Db      241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
QY      301 CTGAAAGAACCTGCGAGAGGCGAAACCGGTGAGAGGCTGCGACCTATGATTTTGTGACAC 360
Db      301 CTGAAAGAACCTGCGAGAGGCGAAACCGGTGAGAGGCTGCGACCTATGATTTTGTGACAC 360
QY      361 TCAAGTTTACAGAGACCAAGGTGATCACTCTGCGACGAGTGTCTGTTTGAAGGATC 420
Db      361 TCAAGTTTACAGAGACCAAGGTGATCACTCTGCGACGAGTGTCTGTTTGAAGGATC 420
QY      421 TTGGTGTCTTACACGCGAGAGATCCGCGACATGTTCCACTGCGCTCTTGTGAGACAC 480
Db      421 TTGGTGTCTTACACGCGAGAGATCCGCGACATGTTCCACTGCGCTCTTGTGAGACAC 480
QY      481 GACTCCGACCTGAGGCTGTCTCGAAGAGTCTCCGCGACGTCGCGCGAGAGGAGACCTG 540
Db      481 GACTCCGACCTGAGGCTGTCTCGAAGAGTCTCCGCGACGTCGCGCGAGAGGAGACCTG 540
QY      541 GAGCAGATTCGAGCAGTACACACCTTCTGTAAGCCGCGCTTCAGAGATTCGCTG 600
Db      541 GAGCAGATTCGAGCAGTACACACCTTCTGTAAGCCGCGCTTCAGAGATTCGCTG 600
QY      601 CCGACAAAGAGTATGCGGATGTATCATCCACAGAGTGAACAATATGTTGCCATC 660
Db      601 CCGACAAAGAGTATGCGGATGTATCATCCACAGAGTGAACAATATGTTGCCATC 660
QY      661 AACCTGATTCGAGCAGTACATCCAGAGATTCGTAAGTGTGACATCTGCAAAATGCGACCGA 720
Db      661 AACCTGATTCGAGCAGTACATCCAGAGATTCGTAAGTGTGACATCTGCAAAATGCGACCGA 720
QY      721 GAGAGGTTCGAATGAGCGAGCTACAGAGCGGACCTTTCTGAGCGAGGAGACCACTGAGG 780
Db      721 GAGAGGTTCGAATGAGCGAGCTACAGAGCGGACCTTTCTGAGCGAGGAGACCACTGAGG 780
QY      781 ATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACCCCACTGA 834
Db      781 ATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACCCCACTGA 834

RESULT 2
US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhimei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/486,725

```

```

; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_files Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (95) .. (928)
; US-10-037-270-546

```

```

Query Match      100.0%; Score 834; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred No. 1,1e-244;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
Db      95 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 154
QY      61 CAGCGCCCTTCTCTGATTAAGGCGTGAAGCGCGGCACTGCGAGCGGAAATCGACCTGTG 120
Db      155 CAGCGCCCTTCTCTGATTAAGGCGTGAAGCGCGGCACTGCGAGCGGAAATCGACCTGTG 214
QY      121 GAGAAGTATGAGGTTCGCGGACAGAAAGAGTGAACAGCGGCGAGGAGTGTG 180
Db      215 GAGAAGTATGAGGTTCGCGGACAGAAAGAGTGAACAGCGGCGAGGAGTGTG 274
QY      181 ATCTGAGCCAGACAGTTCACAGATGCTTGAACGCGACAGACGAGGCGCAAGGCTTG 240
Db      275 ATCTGAGCCAGACAGTTCACAGATGCTTGAACGCGACAGACGAGGCGCAAGGCTTG 334
QY      241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 300
Db      335 AAAGACAGTACAAATTTTGAACATCCAGATGCTTTGATAATGATTTGATGACAGAGCT 394
QY      301 CTGAAAGAACCTGCGAGGCGAAGACCGGTGAGAGTGCAGACCTATGATTTTGTGACAC 360
Db      395 CTGAAAGAACCTGCGAGGCGAAGACCGGTGAGAGTGCAGACCTATGATTTTGTGACAC 454
QY      361 TCAAGTTTACAGAGACCAAGGTGATCTACCTGCGAGAGTGTCTGTTTGAAGGACATC 420
Db      455 TCAAGTTTACAGAGACCAAGGTGATCTACCTGCGAGAGTGTCTGTTTGAAGGACATC 514
QY      421 TTGGTGTCTTACACGCGAGAGATCCGCGACATGTTCCACTGCGCTCTTGTGAGACAC 480
Db      515 TTGGTGTCTTACACGCGAGAGATCCGCGACATGTTCCACTGCGCTCTTGTGAGACAC 574
QY      481 GACTCCGACCTGAGGCTGTCTCGAAGAGTCTCCGCGACGTCGCGCGAGAGGAGACCTG 540
Db      575 GACTCCGACCTGAGGCTGTCTCGAAGAGTCTCCGCGACGTCGCGCGAGAGGAGACCTG 634
QY      541 GAGCAGATTCGAGCAGTACACACCTTCTGTAAGCCGCGCTTCAGAGATTCGCTG 600
Db      635 GAGCAGATTCGAGCAGTACACACCTTCTGTAAGCCGCGCTTCAGAGATTCGCTG 694
QY      601 CCGACAAAGAGTATGCGGATGTATCATCCAGAGAGTGAACAATATGTTGCCATC 660
Db      695 CCGACAAAGAGTATGCGGATGTATCATCCAGAGAGTGAACAATATGTTGCCATC 754
QY      661 AACCTGATTCGAGCAGTACATCCAGAGATTCGTAAGTGTGACATCTGCAAAATGCGACCGA 720
Db      755 AACCTGATTCGAGCAGTACATCCAGAGATTCGTAAGTGTGACATCTGCAAAATGCGACCGA 814
QY      721 GAGAGGTTCGAATGAGCGAGCTACAGAGCGGACCTTTCTGAGCGAGGAGACCACTGAGG 780
Db      815 GAGAGGTTCGAATGAGCGAGCTACAGAGCGGACCTTTCTGAGCGAGGAGACCACTGAGG 874
QY      781 ATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACCCCACTGA 834
Db      875 ATGCTGACCTCTGCGAAACGGTCAATTTGAGTCCAGACAGACCCCACTGA 928

```

```
RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(927)
US-09-896-522-1

Query Match          100.0%; Score 834; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 1.3e-244;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGTCGCGAC 60
DB 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGTCGCGAC 153
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGCGAAGTGCCTGTGT 120
DB 154 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGCGAAGTGCCTGTGT 213
QY 121 GAGAAGATCATGAGTGTCTGAGCAGAACGAGGTGAGACGCGCGAGAGTGTGTC 180
DB 214 GAGAAGATCATGAGTGTCTGAGCAGAACGAGGTGAGACGCGCGAGAGTGTGTC 273
QY 181 ATCTTGAGCCAGGACAGGATTTACAAAGTCTGACGCGAGAGCAGAAAGCCCTTG 240
DB 274 ATCTTGAGCCAGGACAGGATTTACAAAGTCTGACGCGAGAGCAGAAAGCCCTTG 333
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 300
DB 334 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 393
QY 301 CTGAAGAACATCGTGAAGGCGAAACGAGTGAAGTGCAGCTATGATTTTGAACACAC 360
DB 394 CTGAAGAACATCGTGAAGGCGAAACGAGTGAAGTGCAGCTATGATTTTGAACACAC 453
QY 361 TCAAGTTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 420
DB 454 TCAAGTTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 513
QY 421 TTGTTGTTTACAGCCGAGAGATCCGAGATGTTTCAACTGCGGCTCTTCTGAGACAC 480
DB 514 TTGTTGTTTACAGCCGAGAGATCCGAGATGTTTCAACTGCGGCTCTTCTGAGACAC 573
QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
DB 574 GACTCCGAGCTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTG 633
QY 541 GAGCAGATTTCTGACGAGTACACCACTTGTGAAAGCGGCTTCTGAGAGTTCTGCTG 600
DB 634 GAGCAGATTTCTGACGAGTACACCACTTGTGAAAGCGGCTTCTGAGAGTTCTGCTG 693
QY 601 CCGAACAAGAAATGCGGATGATGATATCCCAAGAGAGTGAACAATATGTTGCTCATC 660
DB 694 CCGAACAAGAAATGCGGATGATGATATCCCAAGAGAGTGAACAATATGTTGCTCATC 753
QY 661 AACCTGATCGTAGACATCCAGGACATTTGAAATGTGACATCTGCAATGGACACGA 720
```

```
|||||
DB 754 AACCTGATCGTAGACATCCAGGACATTTGAAATGTGACATCTGCAATGGACACGA 813
QY 721 GAGGTTCAATGCGGAGAGTACAAAGGACCTTTTCTGAGCAGAGGAGACACCTGGG 780
DB 814 GAGGTTCAATGCGGAGAGTACAAAGGACCTTTTCTGAGCAGAGGAGACACCTGGG 873
QY 781 ATGCTGACCTCTGCGAAGAGTCAATTTGAGTCCAGAGAGACCCCACTGA 834
DB 874 ATGCTGACCTCTGCGAAGAGTCAATTTGAGTCCAGAGAGACCCCACTGA 927

RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Query Match          98.7%; Score 823; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 2.9e-241;
Matches 834; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
DB 84 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 143
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAGTGCACCTGTGT 120
DB 144 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGAGTGCACCTGTGT 203
QY 121 GAGAAGATCATGAGTGTCTGAGGACAGAACGAGTGAACAGCGGACGCGAAGGTGTCT 180
DB 204 GAGAAGATCATGAGTGTCTGAGGACAGAACGAGTGAACAGCGGACGCGAAGGTGTCT 263
QY 181 ATCTTGAGCCAGGACAGGATTTTCAAGATCTGACGCGCAGAGAGGCGCAAGGCTTG 240
DB 264 ATCTTGAGCCAGGACAGGATTTTCAAGATCTGACGCGCAGAGAGGCGCAAGGCTTG 323
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 300
DB 324 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 383
QY 301 CTGAAGAACATCGTGAAGGCGAAACGAGTGAAGTGCAGCTATGATTTTGAACACAC 360
DB 384 CTGAAGAACATCGTGAAGGCGAAACGAGTGAAGTGCAGCTATGATTTTGAACACAC 443
QY 361 TCAAGTTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 420
DB 444 TCAAGTTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTGTTTGAAGGCTATC 503
QY 421 TTGTTGTTTACAGCCGAGAGATCCGAGATGTTTCAACTGCGGCTCTTCTGAGACAC 480
DB 504 TTGTTGTTTACAGCCGAGAGATCCGAGATGTTTCAACTGCGGCTCTTCTGAGACAC 563
QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
```



```
Db 526 CCAGTTCACCTTTGACCAACCGGATGCTTTGACAAATGAATCTATTCACAAAACACTCAA 585
Qy 306 GAACATCTGTGAGGGGCAAAAACGCTGAGAGTCCGACCTATGATTTGTGACACACTCAAG 365
Db 586 AGAATACCTGAAAGGAAAAACATCCAGATCCCGGTATGACTTTGTCTCCATTCGCG 645
Qy 366 GTTACCAAGAACCAACGGTGGTCTACCTCGGAGAGTGGTCTGTTGAGGGCAATCTTGGT 425
Db 646 GAAGAGAGAGACAGTACTGTCTATCCCGACAGAGTGGTCTCTTTGAGAGGATCTGGC 705
Qy 426 GTTTCACAGCAGAGAGATCCGGAGCATGTTCCACCTGCGCTCTTCTGTGACACCGACTC 485
Db 706 CTCTCTACCTCCAGAGAGTACAGAGCTGTTCCAGATGAAGCTTTTGTGATACAGATGC 765
Qy 486 CGACGTCAAGCTCTCTCGAAGAGTTCGCGGAGAGTGGC---CGAGGAGGAGCTTGA 542
Db 766 GGAACACCCGGCTCCAGCAGAGATTTAAGGACATCAGCAGAGAGGAGGAGATCTTGA 825
Qy 543 GCAGATTCGTGACGCACTACCAACCTTGTGAAAGCCGGCTTGGAGAGTTCGTGCTGCC 602
Db 826 GCGATTTTATCTCAGATCACTTACGTTCTGTCAGAGCTGCTTGGAGAAATTCGTGGCC 885
Qy 603 GACAAAGAGATATGCCGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATCAA 662
Db 886 AACAAAGAGATATGCTGATGTATCATCCCTAGAGGTCAGATATCTGGTGGCCATCAA 945
Qy 663 CCGTATCTGTGACACATCCAGACATTTGTAATGTGTGACATCTGCAATGGCA 716
Db 946 CCGTATCTGTGACACATCCAGACATCTGTAATGAGGGCCCTCCAAACGGCA 999
```

RESULT 7

```
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIORITY APPLICATION NUMBER: PCT/US00/05888
; PRIORITY FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-300-220
```

Query Match 43.4%; Score 361.6; DB 10; Length 1310;

Best Local Similarity 74.0%; Pred. No. 4.1e-100; Indels 4; Gaps 2;

Matches 484; Conservative 1; Mismatches 165;

```
Qy 66 GCCCTTCCTGATAGGGGTGAGCGGCGACATGCCAGCGGGAAGTCAGCGTGTGTGAAA 125
Db 288 GCCCTTCCTATAGGCGTCA-CGGGGGAACAGCTAGGCGCAAGCTTCCGTGTGTCTAA 346
Qy 126 GATCATGAGTGTCTGTGGACAGAACGAGTGGAAACAGCGCAGCGAAGGTGTCTCT 185
Db 347 GATCTGTGACCTCTGTGGGACAGATGAGTGGACTATCTGCCAGAAAGGTGTCTCT 406
Qy 186 GAGCAGAGAGAGTCTTACAAGGCTCGACGAGACAGACAGAGCCCAAGGCTTTGAAAG 245
Db 407 GAGCCAGAGATAGCTTCTTACCGTGTCTTACCTCGAGACAGAGGCCCAAGCCCTGAA 466
Qy 246 ACAGTACAAATTTTGAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACTTGA 305
```

```
Db 467 CCAGTTCACCTTTGACCAACCGGATGCTTTGACAAATGAATCTATTCACAAAACACTCAA 526
Qy 306 GAACATCTGTGAGGGGCAAAAACGCTGAGAGTCCGACCTATGATTTGTGACACACTCAAG 365
Db 527 AGAATACCTGAAAGGAAAAACATCCAGATCCCGGTATGACTTTGTCTCCATTCGCG 586
Qy 366 GTTACCAAGAACCAACGGTGGTCTACCTCGGAGAGTGGTCTGTTGAGGGCAATCTTGGT 425
Db 587 GAAGAGAGAGACAGTACTGTCTATCCCGACAGAGTGGTCTCTTTGAGAGGATCTGGC 646
Qy 426 GTTTCACAGCAGAGAGATCCGGAGCATGTTCCACCTGCGCTCTTCTGTGACACCGACTC 485
Db 647 CTCTCTACCTCCAGAGAGTACAGAGCTGTTCCAGATGAAGCTTTTGTGATACAGATGC 706
Qy 486 CGACGTCAAGCTCTCTCGAAGAGTTCGCGGAGAGTGGC---CGAGGAGGAGCTTGA 542
Db 707 GGAACACCCGGCTCCAGCAGAGATTTAAGGACATCAGCAGAGAGGAGGAGGAGATCTTGA 766
Qy 543 GCAGATTCGTGACGCACTACCAACCTTGTGAAAGCCGGCTTGGAGAGTTCGTGCTGCC 602
Db 767 GCGATTTTATCTCAGATCACTTACGTTCTGTCAGAGCTGCTTGGAGAAATTCGTGGCC 826
Qy 603 GACAAAGAGATATGCCGATGTATCATCCACGAGAGTGGACAAATATGTTGCCATCAA 662
Db 827 AACAAAGAGATATGCTGATGTATCATCCCTAGAGGTCAGATATCTGGTGGCCATCAA 886
Qy 663 CCGTATCTGTGACACATCCAGACATTTGTAATGTGTGACATCTGCAATGGCA 716
Db 887 CCGTATCTGTGACACATCCAGACATCTGTAATGAGGGCCCTCCAAACGGCA 940
```

RESULT 8

```
US-09-918-995-24042
; Sequence 24042, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIORITY APPLICATION NUMBER: US/09/235,076
; PRIORITY FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24042
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-24042
```

Query Match 26.5%; Score 221; DB 11; Length 472;

Best Local Similarity 73.0%; Pred. No. 3e-57;

Matches 284; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```
Qy 66 GCCCTTCCTGATAGGGGTGAGCGGCGACATGCCAGCGGGAAGTCAGCGTGTGTGAAA 125
Db 84 GCCCTTCCTTATAGGCGTCA-CGGGGGAACAGCTAGGCGCAAGCTTCCGTGTGTCTAA 143
Qy 126 GATCATGAGTGTCTGTGGACAGAACGAGTGGAAACAGCGCAGCGAAGGTGTCTCT 185
Db 144 GATCTGTGACCTCTGTGGGACAGATGAGTGGACTATCTGCCAGAAAGCAAGTGTCTCT 203
Qy 186 GAGCAGAGAGAGTCTTACAAGGCTCGACGAGACAGACAGAGCCCAAGGCTTTGAAAG 245
Db 204 GAGCCAGAGATAGCTTCTTACCGTGTCTTACCTCGAGACAGAGGCCCAAGCCCTGAA 263
Qy 246 ACAGTACAAATTTTGAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACTTGA 305
```

Accession	Sequence	Position
Db	CCAGTTCACCTTTTGACCCACCCGGATGCCCTTTGACATGAATCATTTCCAAACAATCAA	3233
Db	264	
Qy	GAACATCGTGAGAGGCAAAACGGTGGAGGTGGCGACATGATTTTGGACACACTGAAG	365
Db	324	
Db	AGAAATCATCTGAAGGGAAAAACAGTCAGATCCCGCTGATGACTTTTGTCCCAATTCGG	3833
Qy	GTTAACGAGACCAACGGTGTCTAACCCCTCGGACGATGGTCTGTTTGAGGSCATCTTGGT	425
Db	384	
Db	GAGGAGGAGACAGTTACTGTCTATCCGCGACAGCGTGGTCTTTTGAAGGATCTGGC	4433
Qy	GTCTACAGCCAGAGATCCGGGACATG	454
Db	444	
Db	CTTCTACTCCAGAGAGTACGAACTGT	472

```

RESULT 9
US-10-066-543-1843
/ Sequence 1843, Application US/10066543
/ Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indrias, Carol Yoseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
OF COLON CANCER
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066,543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1843
LENGTH: 447
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-1843

```

Query Match	25.8%	Score 214.8;	DB 14;	Length 447;
Best Local Similarity	73.0%;	Pred. No. 2.3e-55;		
Matches 276;	Conservative 0;	Mismatches 102;	Indels 0;	Gaps 0;

QY	66	GGCCCTTCTGATATAGGGGTGATGCGCGGACATCTCCACGCGGAAAGTCGACCGCTGTGTGAA	125
Db	70	GGCCCTTCTTATATAGGGGTGATGCGCGGAAACACTTAGCGGAAGCTTCCGTGTGTCTAA	129
QY	126	GATCATGGAATGCTCTGGACAGACGAGGTGAAACAGGGGACAGGGGAGGTCACTCT	185
Db	130	GATCTGTCACTCTCTGGGCGAATATAGGTGACTATGCCAGAAAGAGGTGCATCTCT	189
QY	186	GAGCCAGAGACAGTTCTTACAGATCCTTGACGCGACAGACAGAGGCCAAGGCTTTGAAAGG	245
Db	190	GAGCCAGATAGTTCTTACCGTGTCTTAACTCTGGAGACAGAGGCCAAGGCCCTGAAAGG	249
QY	246	ACAGTACAAATTTTGAACCATTCAGATGCCCTTGATATATGATTTGACACAGACTCGAA	305
Db	250	CCAGTTCAACTTTGACCACTCCGGAATGCCTTTGACATATCACTCATTTCTCAAAACACTCA	309
QY	306	GAAACATGTGGAGGGCAAAACGGTGAAGTGCACCTATGATTTTGTGACACCTCAG	365
Db	310	AGAAATCACTGAAGGGGAAAAACAGTCAGATCCCGGTATGACTTTGTCTCCACTTCCG	369
QY	366	GTTACCAAGACCAACGCTGTCTACCTCTGCCGACGTGTCTCTTTGAGGCACTTTGGT	425
Db	370	GAAAGAGAGACAGTTACTGTCTATCTCCGACACGTGTGTCTTTGAAAGATCTCGG	429

Qy 426 GTTCTACAGCCAGAGAT 443
 ||||| |||||
 Db 430 CTTCTACTCCAGAGGT 447

```

RESULT 10
US-09-918-995-23923
? Sequence 23923, Application US/09918995
? Publication No. US20030073623A1
? GENERAL INFORMATION:
?   APPLICANT: Hyspec, Inc.
?   TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
?   TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
?   FILE REFERENCE: 20411-756
?   CURRENT APPLICATION NUMBER: US/09/918, 995
?   CURRENT FILING DATE: 2001-07-30
?   PRIOR APPLICATION NUMBER: US/09/235, 076
?   PRIOR FILING DATE: 1999-01-20
?   NUMBER OF SEQ. ID NOS: 18054
?   SOFTWARE: FASTSEQ for Windows Version 3.0
?   SEQ ID NO 23923
?   LENGTH: 455
?   TYPE: DNA
?   ORGANISM: Homo sapiens
?   FEATURE:
?   NAME/KEY: misc_feature
?   LOCATION: (1)...(455)
?   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

```

Query Match	25.2%;	Score 210.4;	DB 11;	Length 455;
Best Local Similarity	72.8%;	Pred. No. 5.1e-54;		
Matches 271; Conservative	0;	Mismatches 101;	Indels 0;	Gaps 0;

OY	66	GCCTTCCTGATAGGGGAGCGCGGACCTGGCAGCGGGAAAGTCGACCGGTGTGGAA	125
Db	84	GCCCTCTCTTAATAGCGCTCAGCGGGGGAAACAGCTAGCGGCAAGCTCTTCGTGTGCTPA	143
OY	126	GATCATGAGTGTGCTGGGACAGAAACGAGGTGGAAACAGCGGACGCGAAAGTGTATCCT	185
Db	144	GATGTGCAAGTCTCTGGGGGCGAAATGAAGTGAATATGTGCCAGAAAGCAGGTGTATCCT	203
OY	186	GAGCCAGGACAGGTTCTACAAAGTCTCTGACGACGAGCAGAAAGGCCAAGCCTTGA	245
Db	204	GAGCAGGATGTGCTTCTACCGGTGCTCTTAACCTCGGAGAGAAAGGCCAAGCCCTGAAGG	263
OY	246	ACAGTACATTTTGAACCATCCAGATGCTTTGATATATGATTTGATGCACAGATCTTGA	305
Db	264	CCAGTTCAACTTTTACCAACCCGAGATGCTTTGAACATTAACCTCATTTCCAAAACCTCAA	323
OY	306	GAACATCGTGGAGGGCAAAACGATGGAGGTGCCGACCATATATTTTGTGACACACTCAAG	365
Db	324	AGAAATCACTAAGGGAACACAGTCCAGATCCCCGTGTATATCTTTGTCTCCCATTTCCG	383
OY	366	GTTACAGAGACCAAGTGTGTCTACCTTCGCGAAGTGTGTTCTGTTGAAGGCATCTTGT	425
Db	384	GAAGAGAGACCAATTACTGTCTATCCCGCAGAGTGTGCTCTTTGAAGGATCTGTGC	443
OY	426	GTTCTACAGCCA	437
Db	444	CTTCTACTCCCA	455

```

RESULT 11
US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US2003019470A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GENOTYPING
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

```



```

FILE REFERENCE: AEOmica-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 15883
LENGTH: 187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: EST HUMAN HIT: A199217.1, EVALUE 1.00e-100
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
OTHER INFORMATION: NT HIT: g14783235, EVALUE 1.00e-100
US-10-029-386-15883

```

```

Query Match          21.9%; Score 183; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 8.8e-46;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      652 GTTGCCATCAACCTGATGTGCGACACATCCAGACATTTCTGAATGTTGACATCTGCATA 711
Db       3 GTTCCCATCAACCTGATGTGCGACACATCCAGACATTTCTGAATGTTGACATCTGCATA 62

QY      712 TGGCACCAGAGAGGTTCCATGCGGCGAGCTACAAGCGGACCTTTCTGAGCCAGGGGAC 771
Db       63 TGGCACCAGAGAGGTTCCATGCGGCGAGCTACAAGCGGACCTTTCTGAGCCAGGGGAC 122

QY      772 CACCTTGAGATGCTGACCTCTTGCCAAACGGTCAACATTTGAGTCCAGACGAGACCCAC 831
Db      123 CACCTTGAGATGCTGACCTCTTGCCAAACGGTCAACATTTGAGTCCAGACGAGACCCAC 182

QY      832 TGA 834
Db      183 TGA 185

RESULT 12
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2183
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7

```

```

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: g11899252, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: B664526.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 2.00e-25
US-10-029-386-2183

```

QY	652	TTTGCAATCAACCTGATGATGACACATATCCAGGACATTGGAATGGGACATCTGCAAA	711
Db	26	GTTCGCATCAACCTGATGATGACACATATCCAGGACATTGGAATGGGACATCTGCAAA	85
QY	712	TGGCACCAGGAGGGATCCCATG3GGCGAGCTTACAAAGCGGACCTTTCTGAGCCAGGGAC	771
Db	86	TGGCACCAGGAGGGATCCCATG3GGCGAGCTTACAAAGCGGACCTTTCTGAGCCAGGGAC	145
QY	772	CACCCCTGGGATGCTGACCTCTGGCAAAAGGTCACTTTGGAGTCCAGCAGCAGACCCAC	831
Db	146	CACCCCTGGGATGCTGACCTCTGGCAAAAGGTCACTTTGGAGTCCAGCAGCAGACCCAC	205
QY	832	TGA	834
Db	206	TGA	208

```

RESULT 13
US-09-764-877-2804
; Sequence 2804, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2804
; LENGTH: 9732
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2804

Query Match      21.9%; Score 183; DB 10; Length 9732;
Best Local Similarity 100.0%; Pred. No. 4e-45;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      652  GTTCCATTCACACCTGATCGTGCAGACACATCCAGGACATTCGTAATGGTGCATCTGCAAA 711
Db      3901  GTTGCCATTCACACTGATCGTGCAGACACATCCAGGACATTCGTAATGGTGCATCTGCAAA 3966

Oy      712  TGGACACCGAGAGGGTCCCATGGGCGGAGCTACAAGCGGACCTTTTCTGAGCCAGGGGAC 771
Db      3961  TGGACACCGAGAGGGTCCCATGGGCGGAGCTACAAGCGGACCTTTTCTGAGCCAGGGGAC 4020

Oy      772  CACCTTGGGATGCTGACACTCTGCGCAACGGTGCACATTTGAGTGCAGAGACAGACCCAC 831
Db      4021  CACCTTGGGATGCTGACACTCTGCGCAACGGTGCACATTTGAGTGCAGAGACAGACCCAC 4080

Oy      832  TGA 834
Db      4081  TGA 4083

RESULT 14
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125.
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2805

Query Match          21.9%; Score 183; DB 10; Length 19125;
Best Local Similarity 100.0%; Pred. No. 5,2e-45;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 652 GTTGCCATCAACCTGATGTGACGACATCCAGACATTTGAAATGGTACATCTGCAAA 711
DB 5997 GTTGCCATCAACCTGATGTGACGACATCCAGACATTTGAAATGGTACATCTGCAAA 6056
QY 712 TGGGACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGAC 771
DB 6057 TGGGACCGAGAGGGTCCAAATGGGCGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGAC 6116
QY 772 CACCCCTGGATGTGACCTCTGGCAAAAGGTCACTTGGAGTGCAGAGAGACCCAC 831
DB 6117 CACCCCTGGATGTGACCTCTGGCAAAAGGTCACTTGGAGTGCAGAGAGACCCAC 6176
QY 832 TGA 834
DB 6177 TGA 6179

RESULT 15
US-10-251-186-14
; Sequence 14, Application US/10251186
; Publication No. US20030180745A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Y. Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030180745A1 Nucleic Acids and
; FILE REFERENCE: 789CIP2CDV1
; CURRENT APPLICATION NUMBER: US/10/251,186
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 14
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(1909)
US-10-251-186-14

Query Match          21.1%; Score 175.6; DB 12; Length 2058;
Best Local Similarity 58.4%; Pred. No. 4.1e-43;
Matches 326; Conservative 0; Mismatches 229; Indels 3; Gaps 1;
```

```
DB 669 ACAAGTGTGACTGAGCAGACAGAGGACCGGACGACCAACCACTTCACTTGAC 728
QY 263 ATCCAGATGCTTTGTAATGATTTGATGACAGAGACTCTGAAGACATCTGAGAGGCA 322
DB 729 ACCCAGATGCTTTGACTTGACCTTCATTTCCACCTTCAGAGAGCTGAACAGGGGA 788
QY 323 AAACGGTGAAGTGCAGACCTATGATTTTGTACACACTCAAGTTAACAGAGACCGG 382
DB 789 AGAGTGTCAAGGTGCTCCATTTATGACTTACCCAGCAGACCGGAAAGAGACTGAGAAA 848
QY 383 TGGTCTAACCTTGGGAGCGTGGTCTGTTGAGGGCATCTTGTGTTCTACAGCCAGAGA 442
DB 849 CACTGTATGTGTCAAAAGTCATCATCTTTGAGGGCATATGACCTTTGCTGACAAAGCAC 908
QY 443 TCCGGGACATGTTCCACCTGCGCTCTTGTGTGACACCACTCCGACGTACAGCTGTCTC 502
DB 909 TGTGGAGCTCTGAGCATGAAATCTTTGTGACACAGCTCCGACATCCGCTGTGAC 968
QY 503 GAAAGTTCCTCGGAGCGTGC---CCGAGGAGGGGACCTGAGCAGATTCTGACGAGT 559
DB 969 GCGGCTGCGCCGGACATGATGAGCGCGCGGACATCGAGGTGTATCAAGCAGT 1028
QY 560 ACAACAACCTTGTGAAGCCGCTTCGAGAGTTCCTGCGGACAAAAGATATGCCG 619
DB 1029 ACAACAAGTTTGTCAAGCCCTCTTCGACCAAGTACATCCAGCCCAATGCGCTGGCAG 1088
QY 620 ATGTGATCATCCACAGAGAGTGAACATATGTGTCCATCAACTGATGTGACAGCA 679
DB 1089 ACATCGTGTGCTCCAGAGGGAGCGGACACGCTGSCCATTCGACTGATTTGTGACAGCG 1148
QY 680 TCCAGGACATTTGGAATG 697
DB 1149 TGCACAGCCAGCTGAGG 1166
```

Search completed: November 25, 2003, 02:19:00
Job time : 307.209 secs


```

Db      181 ATCTGAGCCGAGACAGTTCTACAGTCTCTGAAGCGGAGAGCAAGGCGCTTG 240
Qy      241 AAAGACAGTACAAATTTTGCACATCCAGATCCCTTTGTAATGATTTGATGCACAGACT 300
Db      241 AAAGACAGTACAAATTTTGCACATCCAGATCCCTTTGTAATGATTTGATGCACAGACT 300
Qy      301 CTGAAGAACATCGTGGAGGGGCAAAACGGTGGAGTGGCGGACCTTGTGATTTGTAACAC 360
Db      301 CTGAAGAACATCGTGGAGGGGCAAAACGGTGGAGTGGCGGACCTTGTGATTTGTAACAC 360
Qy      361 TCAAGTTACAGAGACACCGTGGTCTACCTCGGAGCGTGTCTGTGTTGAGGGGATC 420
Db      361 TCAAGTTACAGAGACACCGTGGTCTACCTCGGAGCGTGTCTGTGTTGAGGGGATC 420
Qy      421 TTGGTGTCTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
Db      421 TTGGTGTCTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
Qy      481 GACTCCGACGTCAGGCTGTCTGAAAGTTCCTCGGGAGCGGCGGAGGGGAGGACCTG 540
Db      481 GACTCCGACGTCAGGCTGTCTGAAAGTTCCTCGGGAGCGGCGGAGGGGAGGACCTG 540
Qy      541 GAGCAGATTCTGACGACAGTACACACCTTCTGTGAAAGCGGCTTCTGAGAGTTCTGCTG 600
Db      541 GAGCAGATTCTGACGACAGTACACACCTTCTGTGAAAGCGGCTTCTGAGAGTTCTGCTG 600
Qy      601 CCGACAAAGAGTATGCCAGATGATCATCCCAAGAGAGTGGACATATGTTGCCATC 660
Db      601 CCGACAAAGAGTATGCCAGATGATCATCCCAAGAGAGTGGACATATGTTGCCATC 660
Qy      661 AACCTGATCGTGCAGACATCCAGACATCTCTGATGATGTCATCTGCAAAATGGCACCGA 720
Db      661 AACCTGATCGTGCAGACATCCAGACATCTCTGATGATGTCATCTGCAAAATGGCACCGA 720
Qy      721 GGAAGGTTCATATGGGCGGAGCTAACAAGCGACCTTTCTGAGCGGAGGACCACTTGGG 780
Db      721 GGAAGGTTCATATGGGCGGAGCTAACAAGCGACCTTTCTGAGCGGAGGACCACTTGGG 780
Qy      781 ATGCTGACCTTGGCAACGGTCACTTTGGAGTCCAGACAGACCCCACTGA 834
Db      781 ATGCTGACCTTGGCAACGGTCACTTTGGAGTCCAGACAGACCCCACTGA 834

```

RESULT 2

```

US-10-037-270-546
; Sequence 546, Application US/10037270
; Publication No. US20030104529a1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Zhimei
; APPLICANT: Fillinghaas, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529a1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725

```

```

; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_files Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-10-037-270-546

```

```

Query Match      100.0% Score 834; DB 14; Length 1288;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
Db      95 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 154
Qy      61 CAGCGGCTCTTCTGATAGGGGTGACCGCGGCACTGCGAGCGGGAAGTCCGCTGTGT 120
Db      155 CAGCGGCTCTTCTGATAGGGGTGACCGCGGCACTGCGAGCGGGAAGTCCGCTGTGT 214
Qy      121 GAGAAATCATGAGATTGCTGGGACAGAAAGAGTGGAAACAGCGGCGGAGAGTGTGTC 180
Db      215 GAGAAATCATGAGATTGCTGGGACAGAAAGAGTGGAAACAGCGGCGGAGAGTGTGTC 274
Qy      181 ATCTGAGCCGAGACAGTCTTCAAGAGTCTTACGCGGACAGAGCAAGGCGAAAGCTTTG 240
Db      275 ATCTGAGCCGAGACAGTCTTCAAGAGTCTTACGCGGACAGAGCAAGGCGAAAGCTTTG 334
Qy      241 AAAGACAGTACAAATTTTGCACATCCAGATGCTTGTATATGATTTGATGCAAGACT 300
Db      335 AAAGACAGTACAAATTTTGCACATCCAGATGCTTGTATATGATTTGATGCAAGACT 394
Qy      301 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGTGGCGGACCTATGATTTTGTGACAC 360
Db      395 CTGAAGAACATCGTGGAGGGCAAAACGGTGGAGTGGCGGACCTATGATTTTGTGACAC 454
Qy      361 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGACGTTGTTGAGGGGATC 420
Db      455 TCAAGTTACAGAGACCAAGGTGTCTACCTGCGGACGTTGTTGAGGGGATC 514
Qy      421 TTGGTGTCTACAGCAGAGAGATCCGGGACATTTCCACTGGGCTCTTCTGTGACACC 480
Db      515 TTGGTGTCTACAGCAGAGAGATCCGGGACATTTCCACTGGGCTCTTCTGTGACACC 574
Qy      481 GACTCCGACGTCAGGCTGTCTGAAAGTTCCTCGGAGCGTGGCGGAGGAGGACTTG 540
Db      575 GACTCCGACGTCAGGCTGTCTGAAAGTTCCTCGGAGCGTGGCGGAGGAGGACTTG 634
Qy      541 GAGCAGATTCTGACGAGTACACACCTTCTGTAAGCGGCGCTTCTGAGAGATTCTGCTTG 600
Db      635 GAGCAGATTCTGACGAGTACACACCTTCTGTAAGCGGCGCTTCTGAGAGATTCTGCTTG 694
Qy      601 CCGACAAAGAGTATGCCAGTGTGATCCCAAGAGAGTGGACAAATATGTTGGCATC 660
Db      695 CCGACAAAGAGTATGCCAGTGTGATCCCAAGAGAGTGGACAAATATGTTGGCATC 754
Qy      661 AACCTGATCGTGCAGACATCCAGACATCTCTGATGATGTCATCTGCAAAATGGCACCGA 720
Db      755 AACCTGATCGTGCAGACATCCAGACATCTCTGATGATGTCATCTGCAAAATGGCACCGA 814
Qy      721 GGAAGGTTCATATGGGCGGAGCTAACAAGCGACCTTTCTGAGCGGAGGACCACTTGGG 780
Db      815 GGAAGGTTCATATGGGCGGAGCTAACAAGCGACCTTTCTGAGCGGAGGACCACTTGGG 874
Qy      781 ATGCTGACCTTGGCAACGGTCACTTTGGAGTCCAGACAGACCCCACTGA 834
Db      875 ATGCTGACCTTGGCAACGGTCACTTTGGAGTCCAGACAGACCCCACTGA 928

```

this hole is good

RESULT 3
US-09-896-522-1
; Sequence 1, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)...(927)
US-09-896-522-1

Query Match 100.0%; Score 834; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
DB 94 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 153
QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGCTCTCCGCGGAGAGTGAACCTGTGT 120
DB 154 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGCTCTCCGCGGAGAGTGAACCTGTGT 213
QY 121 GAGAGATCATGAGTGTCTGAGGACAGAGAGGTGAGAGAGCGCGGAGAGTGTCTG 180
DB 214 GAGAGATCATGAGTGTCTGAGGACAGAGAGGTGAGAGAGCGCGGAGAGTGTCTG 273
QY 181 ATCTGAGCAGGACAGGATCTCAAGAGTCTTCAAGGACGAGGAGGAGGAGGAGG 240
DB 274 ATCTGAGCAGGACAGGATCTCAAGAGTCTTCAAGGAGGAGGAGGAGGAGGAGG 333
QY 241 AAGGAGACGATCAATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 334 AAGGAGACGATCAATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 393
QY 301 CTGAAGACATCGTGAAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 394 CTGAAGACATCGTGAAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTTGTGAGGAGGATC 420
DB 454 TCAAGGTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTTGTGAGGAGGATC 513
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGAGATGTTTCACTGCGGCTTCTGAGACAC 480
DB 514 TTGGTGTCTTACAGCCAGGAGATCCGAGATGTTTCACTGCGGCTTCTGAGACAC 573
QY 481 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGAGGAGTGGGCGAGGAGGAGGAG 540
DB 574 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGAGGAGTGGGCGAGGAGGAGGAG 633
QY 541 GAGCAGATTCTGACGAGTACCACTTCTGAGAGCGGCTTCTGAGAGGATCTGCTG 600
DB 634 GAGCAGATTCTGACGAGTACCACTTCTGAGAGCGGCTTCTGAGAGGATCTGCTG 693
QY 601 CCGAGCAAGAGATGCGGATGATATCCACAGAGAGTGAACAATATGTTGATCCATC 660
DB 694 CCGAGCAAGAGATGCGGATGATATCCACAGAGAGTGAACAATATGTTGATCCATC 753
QY 661 AACCTGATGAGACATCCAGAGATCTGATGATGTGATCAATCTGCAATGAGACCGA 720

DB 754 AACCTGATGAGACATCCAGAGATCTGATGATGTGATCAATCTGCAATGAGACCGA 813
QY 721 GAGAGTCCATGAGGAGGAGGAGTCAAGAGGAGCTTTTCTGAGGAGGAGGAGGAG 780
DB 814 GAGAGTCCATGAGGAGGAGGAGTCAAGAGGAGCTTTTCTGAGGAGGAGGAGGAG 873
QY 781 ATGCTGACCTTGGCAAGGAGGAGGAGTGTGAGTCCAGAGGAGGAGGAGGAGGAG 834
DB 874 ATGCTGACCTTGGCAAGGAGGAGGAGTGTGAGTCCAGAGGAGGAGGAGGAGGAG 927

RESULT 4
US-09-833-381-2048
; Sequence 2048, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: (Robison, Keith E.)
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2048
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1648)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-2048

Query Match 85.6%; Score 714; DB 10; Length 1648;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGAGGAGGAGGAGG 60
DB 84 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGAGGAGGAGGAGG 143
QY 61 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGAGCTGCGAGGAGGAGGAGGAGGAG 120
DB 144 CAGCGGCGCTTCTCTGATAGGGGTGAGCGCGGAGCTGCGAGGAGGAGGAGGAGG 203
QY 121 GAGAGATCATGAGTGTCTGAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 204 GAGAGATCATGAGTGTCTGAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 263
QY 181 ATCTGAGCAGGACAGGATCTCAAGGTCCTGACGCGGAGGAGGAGGAGGAGGAGG 240
DB 264 ATCTGAGCAGGACAGGATCTCAAGGTCCTGACGCGGAGGAGGAGGAGGAGGAGG 323
QY 241 AAGGAGACGATCAATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 300
DB 324 AAGGAGACGATCAATTTTGAACATCCAGATGCTTTATATATGATTTGATGACAGACT 383
QY 301 CTGAAGACATGATGAGGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
DB 384 CTGAAGACATGATGAGGAGGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 443
QY 361 TCAAGGTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTTGTGAGGAGGATC 420
DB 444 TCAAGGTTACAGAGACCAAGTGTCTACCTGCGGACGTGTCTTGTGAGGAGGATC 503
QY 421 TTGGTGTCTTACAGCCAGGAGATCCGAGATGTTTCACTGCGGCTTCTGAGACAC 480
DB 504 TTGGTGTCTTACAGCCAGGAGATCCGAGATGTTTCACTGCGGCTTCTGAGACAC 563
QY 481 GACTCCGAGTCAAGGCTGTCTGAAAGTCTCCGAGGAGTGGGCGAGGAGGAGGAGG 540

614

```
Db 564 GACTCCGACGTCAGGCTCTTCGAAGAGTCTCCGGGACGTCGCCGAGGAGGAGCCTG 623
Qy 541 GAGCAGATTCTGACGACATACCACTTCTGTGAAGCCGCTTGAAGAGTCTGCTG 600
Db 624 GAGCAGATTCTGACGACATACCACTTCTGTGAAGCCGCTTGAAGAGTCTGCTG 683
Qy 601 CCGGAAAGAGTATGCGGATGATATCCGAGAGATGACATATGTTGCTAT 659
Db 684 CCGGAAAGAGTATGCGGATGATATCCGAGAGATGACATATGTTGCTAT 743
Qy 660 CAACCTGATCTGACGACATCCAGACATCTGTAATGTGACATTCGAAATGGACCG 719
Db 744 CAACCTGATCTGACGACATCCAGACATCTGTAATGTGACATTCGAAATGGACCG 803
Qy 720 AGAGGGTCCATAGGGCGGAGCTCAAGCGGACCTTTTCTGAGCCGAGGACCACTTG 779
Db 804 AGAGGGTCCATAGGGCGGAGCTCAAGCGGACCTTTTCTGAGCCGAGGACCACTTG 863
Qy 780 GATGCTGACCTCTGCGAAAGCTGACATTTGAGTCCGAGGACGACCCCACTGA 834
Db 864 GATGCTGACCTCTGCGAAAGCTGACATTTGAGTCCGAGGACGACCCCACTGA 918
```

RESULT 5

```
US-09-918-995-30379
; Sequence 30379, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918, 995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 30379
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30379
```

date is good

```
Query Match 48.3%; Score 403; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 2,4e-207;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 GCCACGGGAGTGAACCGTGTGTGAAGATCATGAGTTGCTGGACAGAGAGTG 156
Db 41 GCCACGGGAGTGAACCGTGTGTGAAGATCATGAGTTGCTGGACAGAGAGTG 100
Qy 157 GAACAGCGGCGAGCGAAGTGTCTATCTGAGCCAGAGCAGTTCTCAAGTCTTGA 216
Db 101 GAACAGCGGCGAGCGAAGTGTCTATCTGAGCCAGAGCAGTTCTCAAGTCTTGA 160
Qy 217 GCAAGAGCAAGGCGCAAGGCTTGAAGAGCAGTACAAATTTGACATCCAGAGCTTT 276
Db 161 GCAAGAGCAAGGCGCAAGGCTTGAAGAGCAGTACAAATTTGACATCCAGAGCTTT 220
Qy 277 GATAATGATTTGATGACAGAGCTCTGAAGAACATCTGTGAGGCGCAAAACGTTGAG 336
Db 221 GATAATGATTTGATGACAGAGCTCTGAAGAACATCTGTGAGGCGCAAAACGTTGAG 280
Qy 337 CCGACCTATGATTTGTGACACACTCAAGGTTACCAAGACCAAGGTTGCTTACCTGCG 396
Db 281 CCGACCTATGATTTGTGACACACTCAAGGTTACCAAGACCAAGGTTGCTTACCTGCG 340
```

```
Qy 397 GAGTGTTCTGTTGAGGGGACATCTTGTTCTACAGCAGGAGATCCGGAGCATGTTG 456
Db 341 GAGTGTTCTGTTGAGGGGACATCTTGTTCTACAGCAGGAGATCCGGAGCATGTTG 400
Qy 457 CACCTGGCTCTTCTGTTGAGCAGCGACTCCGAGCTGAGGCTGT 499
Db 401 CACCTGGCTCTTCTGTTGAGCAGCGACTCCGAGCTGAGGCTGT 443
```

RESULT 6

```
US-10-029-386-15883
; Sequence 15883, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029, 386
; CURRENT FILING DATE: 2001-12-20
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15883
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EST HUMAN HIT: A1992171.1, EVALUE 1.00e-100
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 4.00e-26
; OTHER INFORMATION: NT HIT: g114783235, EVALUE 1.00e-100
US-10-029-386-15883
```

```
Query Match 21.9%; Score 183; DB 12; Length 187;
Best Local Similarity 100.0%; Pred. No. 2,4e-88;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 652 GTTGCAATCAACTGATCGTGCAGACATCCAGACATCTGTAATGTGACATCTGCAAA 711
Db 3 GTTGCAATCAACTGATCGTGCAGACATCCAGACATCTGTAATGTGACATCTGCAAA 62
Qy 712 TGGCACCAGAGAGGTCCTCAATGGGCGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGC 771
Db 63 TGGCACCAGAGAGGTCCTCAATGGGCGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGC 122
Qy 772 CACCTGGATGTGACCTCTGCGAAACGTTCAATTTGAGGTCGAGCAGACAGCCCGAC 831
Db 123 CACCTGGATGTGACCTCTGCGAAACGTTCAATTTGAGGTCGAGCAGACAGCCCGAC 182
Qy 832 TGA 834
Db 183 TGA 185
```

RESULT 7

```
US-10-029-386-2183
; Sequence 2183, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
```

```

1  TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
2  FILE REFERENCE: AEOMICA-X-2
3  CURRENT APPLICATION NUMBER: US/10/029,386
4  CURRENT FILING DATE: 2001-12-20
5  NUMBER OF SEQ ID NOS: 34288
6  SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
7  SEQ ID NO 2183
8  LENGTH: 510
9  TYPE: DNA
10 ORGANISM: Homo sapiens
11 FEATURE:
12 OTHER INFORMATION: MAP TO CHR9.1
13 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
14 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
15 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
16 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
17 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
18 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
19 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
20 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
21 OTHER INFORMATION: NT HIT: G11889252, EVALU0 0.00e+00
22 OTHER INFORMATION: EST HUMAN HIT: BF664526.1, EVALU0 0.00e+00
23 OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU0 2.00e-25
24 US-10-029-386-2183

```

Query Match	21.9%	Score 183;	DB 12;	length 510;
Best Local Similarity	100.0%	Pred. No.	2.3e-88;	
Matches 183;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0

Oy 652 GTTGCCATCAACCTGATCGTGCAGCACATCCAGACATTCTGAATGGTAGCATCTGCCAA 71
|||||
Db 26 GTTGCCATCAACCTGATCGTGCAGCACATCCAGACATTCTGAATGGTAGCATCTGCCAA 85

QY 712 TGGCAACGAGAGGGCTTCCAATGGGCGGAGCTTCAAGGCGACCTTTTCTGAGCCAGGGGAC 7712

Db 86 TGGCAACGAGAGGGCTTCCAATGGGCGGAGCTTCAAGGCGACCTTTTCTGAGCCAGGGGAC 1455

Oy 772 CACCCTGGAGTCTGACCTCTGGCAACGGTCACATTGGAGTCCAGACGACGCCAC 831
 Db 146 CACCCTGGAGTCTGACCTCTGGCAACGGTCACATTGGAGTCCAGACGACGCCAC 205

OY	832	TGA	834
Db	206	TGA	208

```

RESULT 8
US-09-764-877-2804
: Sequence 2804, Application US/09764877
: Patent No. US20020147140A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
: CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 4031
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2804
: LENGTH: 9732
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-877-2804

```

Query Match	21.9%	Score 183	DB 10	Length 9732
Best Local Similarity	100.0%	Pred. No.	2.1e-88	
Matches 183; Conservative	0	Mismatches	0	Gaps 0

Qy	652	711
GTTTCATCAACCGATCGGCAGACATCCAGCAATTGTGAATGTCATTCGCAAA		
Db	3901	3966
GTTTCATCAACCGATCGGCAGACATCCAGCAATTGTGAATGTCATTCGCAAA		

Accession	Sequence	Length
Oy	TGACACGAGAGGGGTCAATTGGCGAGGATCAACAGCGACCTTTCTGAGCCACAGGGAC	771
Db	TGACACGAGAGGGGTCAATTGGCGAGGATCAACAGCGACCTTTCTGAGCCACAGGGAC	4020
Oy	CACCTGGGATGCTGACCTCTCGGAAACGGTCACTTTGAGTCCACAGGACCCAC	831
Db	CACCTGGGATGCTGACCTCTCGGAAACGGTCACTTTGAGTCCACAGGACCCAC	4080

Qy	832	TGA	834
Db	4081	TGA	4083

```

RESULT 9
US-09-764-877-2805
; Sequence 2805, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2805
; LENGTH: 19125
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-877-2805

```

Query Match	21.9%	Score 183	DB 10	Length 19125
Best Local Similarity	100.0%	Pred. No. 2.1e-88		
Matches 183	Conservative 0	Mismatches 0	Indels 0	Gaps 0

0y 652 GTTGGCAGCAACCTGATCGTGGAGCAATCCAGAGACATCTGATGGGAGACATCTGCAA 711
Db 5997 GTTGGCAGCAACCTGATCGTGGAGCAATCCAGAGACATCTGATGGGAGACATCTGCAA 6056

QY 712 TGGCACCAGAGAGGGTCCAAATGGGGCGAGCTCAAGCGGACCCCTTTCTCGAGCCAGGGGAC 771

Db 6057 TGGCACCAGAGAGGGTCCAAATGGGGCGAGCTCAAGCGGACCCCTTTCTCGAGCCAGGGGAC 6116

QY 772 CACCTGTGGATGCTGACCTCTGGCAAAAGTCAATTGGAGTCAGACGACACCCAC 831
Db 6117 CACCTGTGGATGCTGACCTCTGGCAAAAGTCAATTGGAGTCAGACGACACCCAC 6176

Qy	832	TGA	834
Db	6177	TGA	6179

```

RESULT 10
US-09-908-975-24159
: Sequence 24159, Application US/09908975
: Publication No. US20030165843A1
:
: GENERAL INFORMATION:
:
: APPLICANT: SHOSHAN, AVI
: APPLICANT: WASERMAN, ALON
: APPLICANT: MINTZ, ELI
: APPLICANT: MINTZ, ELI
: APPLICANT: FAIGLER, SIMCHON
:
: TITLE OF INVENTION: OLGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE V
:
: TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
:
: FILE REFERENCE: 36688-0005
:
: CURRENT APPLICATION NUMBER: US/09/908,975
:
: CURRENT FILING DATE: 2001-07-20
:
: PRIOR APPLICATION NUMBER: US 60/287,724
:
: PRIOR FILING DATE: 2001-05-02
:
: PRIOR APPLICATION NUMBER: US 60/221,607
:
: PRIOR FILING DATE: 2000-07-28
:
: NUMBER OF SEQ ID NOS: 32337
:

```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24159
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-24159
```

```
Query Match          3.1%; Score 26; DB 12; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY          493 AGGCTGTCTCGAAGATTCTCCGGA 518
              |||||
Db           16 AGGCTGTCTCGAAGATTCTCCGGA 41
```

```
RESULT 11
US-09-925-300-220
; Sequence 220, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-300-220
```

```
Query Match          2.8%; Score 23; DB 10; Length 1310;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY          667 ATCGGCGACATCCAGACAT 689
              |||||
Db           891 ATGTTGCGACATCCAGACAT 913
```

```
RESULT 12
US-10-098-841-53
; Sequence 53, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehtman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1 Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
```

```
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_pl_genes Version 1.0
; SEQ ID NO 53
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (290)..(1075)
US-10-098-841-53
```

```
Query Match          2.8%; Score 23; DB 13; Length 1402;
Best Local Similarity 100.0%; Pred. No. 0.079;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY          667 ATCGTGCACATCCAGACAT 689
              |||||
Db           950 ATCGTGCACATCCAGACAT 972
```

```
RESULT 13
US-10-066-543-1843
; Sequence 1843, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguo
; APPLICANT: Eyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Joseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1843
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1843
```

```
Query Match          2.6%; Score 22; DB 14; Length 447;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY          173 AGGTGTCATCTGAGCAGGA 194
              |||||
Db           177 AGGTGTCATCTGAGCAGGA 198
```

```
RESULT 14
US-09-918-995-23923
; Sequence 23923, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
```



```

; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23923
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-23923

```

```

Query Match      2.6%; Score 22; DB 11; Length 455;
Best Local Similarity 100.0%; Pred.No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      173 AGGTGTCATCTGAGCCAGGA 194
      |||||||||||||||||||
DB      191 AGGTGTCATCTGAGCCAGGA 212

```

```

RESULT 15
US-09-918-995-24042
; Sequence 24042, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hvaeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24042
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(472)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-24042

```

```

Query Match      2.6%; Score 22; DB 11; Length 472;
Best Local Similarity 100.0%; Pred.No. 0.28;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      173 AGGTGTCATCTGAGCCAGGA 194
      |||||||||||||||||||
DB      191 AGGTGTCATCTGAGCCAGGA 212

```

Search completed: November 25, 2003, 07:27:17
 Job time : 675.367 secs

This Page Blank (uspto)

FEATURES
source location/Qualifiers
1. 1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 234 a 311 c 361 g 209 t 3 others
ORIGIN
Query Match 96.5%; Score 805.2; DB 12; Length 1118;
Best Local Similarity 98.9%; Pred. No. 1.3e-188;
Matches 821; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
DB 47 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 106
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCGTGTGT 120
DB 107 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCGTGTGT 166
QY 121 GAGAAATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGCACGCGAAGGTGTC 180
DB 167 GAGAAATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGCACGCGAAGGTGTC 226
QY 181 ATCTGAGCCAGGACAGGTTCTAAGAGTCTGACGCGACGAGAGGCGCAAGGCGCTTG 240
DB 227 ATCTGAGCCAGGACAGGTTCTAAGAGTCTGACGCGACGAGAGGCGCAAGGCGCTTG 286
QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
DB 287 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 346
QY 301 CTGAAGAAATCATGAGGTGAGGCGAAGCGGTGAGGTCGACCTATGATTTTGTGACAC 360
DB 347 CTGAAGAAATCATGAGGTGAGGCGAAGCGGTGAGGTCGACCTATGATTTTGTGACAC 406
QY 361 TCAAGTTACAGAGACCAAGGTGTCTACCTCGGACGAGTGTCTGTTGAGGGATC 420
DB 407 TCAAGTTACAGAGACCAAGGTGTCTACCTCGGACGAGTGTCTGTTGAGGGATC 466
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACC 480
DB 467 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTCTGTGACACC 526
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG 540
DB 527 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCGAGGAGGACCTG 586
QY 541 GAGAGATTCTGAGCGAGTACACGACCTGTGTGAAGCGCGCTTCGAGGAGTTCTGCTG 600
DB 587 GAGAGATTCTGAGCGAGTACACGACCTGTGTGAAGCGCGCTTCGAGGAGTTCTGCTG 646
QY 601 CCGACAAAGAGTATGCGATGTGATATCCACAGAGAGTGAACAATATGTTGCCATC 660
DB 647 CCGACAAAGAGTATGCGATGTGATATCCACAGAGAGTGAACAATATGTTGCCATC 706
QY 661 AACCTGATCTGTCAGCAGCATCAGAGCAATTCGATGAGTGAATTCGCAATGACACCGA 720
DB 707 AACCTGATCTGTCAGCAGCATCAGAGCAATTCGATGAGTGAATTCGCAATGACACCGA 766
QY 721 GAGAGGTTCATGAGCGGAGCTACAAGCGACCTTTCTGAGCGAGGAGCAACCCCTGG- 779
DB 767 GAGAGGTTCATGAGCGGAGCTACAAGCGACCTTTCTGAGCGAGGAGCAACCCCTGGN 826
QY 780 GATGCTACCTCTGCAAAACGGTCACTTTGAGTCCAGCAGACGACCC 829

DB 827 GATGCTACCTCTGCAAAACGGTCACTTTGAGTCCAGAGACGACCC 876
RESULT 2
BX400889 998 bp mRNA linear EST 13-MAY-2003
LOCUS BX400889 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK003YL24 5-PRIME, mRNA sequence.
ACCESSION BX400889
VERSION BX400889.1 GI:30622359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK003YD12Q1&cluster=4968.r Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK003YD12Q1.
FEATURES
source location/Qualifiers
1. 998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK003YL24"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_id="HELA"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 223 a 263 c 314 g 186 t 12 others
ORIGIN
Query Match 93.7%; Score 781.6; DB 13; Length 998;
Best Local Similarity 98.9%; Pred. No. 8.3e-183;
Matches 825; Conservative 2; Mismatches 3; Indels 4; Gaps 4;
QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
DB 67 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 125
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCGTGTGT 120
DB 126 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCGTGTGT 185
QY 121 GAGAAATCATGAGTTGCTGGGACAGAACGAGTGAACAGCGGACGCGAAGGTGTC 180
DB 186 GAGAAATCATGAGTTGCTGGGACAGAACGAGTGAACAGCGGACGCGAAGGTGTC 245
QY 181 ATCTGAGCCAGGACAGGTTCTAAGAGTCTGACGCGACAGACGAGAGGCGCAAGGCTTG 240
DB 246 ATCTGAGCCAGGACAGGTTCTAAGAGTCTGACGCGACAGACGAGAGGCGCAAGGCTTG 305
QY 241 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
DB 306 AAAGACAGTACATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGGACT 365
QY 301 CTGAAGAAATCATGAGGTGAGGCGAAGCGGTGAGGTCGACCTATGATTTTGTGACAC 360

```

Db      366 CTGAAGACATCGTGGAGGGCAAAACGGTGAAGTGCCGACCTATGATTTTGTGACACAC 425
Qy      361 TCAAGTTACAGAGACAGCGTGTCTACCTGCGGACCTGTCTTGTGAGGGGATC 420
Db      426 TCAAGTTACAGAGACAGCGTGTCTACCTGCGGACCTGTCTTGTGAGGGGATC 485
Qy      421 TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGAGACAC 480
Db      486 TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGAGACAC 545
Qy      481 GACTCCGACGTGAGCTGTCTCGAAGATTCTCCGGACGTGCGCGGAGGAGGACCTG 540
Db      546 GACTCCGACGTGAGCTGTCTCGAAGATTCTCCGGACGTGCGCGGAGGAGGAGGACCTG 605
Qy      541 GAGCAATTTGAGAGGATACCACTTGTGAGGCGGCTTTCAGAGAGTTCTGCTG 600
Db      606 GACCAAGTTCTGACGA-TCACACACCTTGTGAGAGCGGCTTTCAGAGAGTTCTGCTG 664
Qy      601 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGTGGACAAATATGTTGGCATC 660
Db      665 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGTGGACAAATATGTTGGCATC 724
Qy      661 AACCTGATCGTGAGCAGCATCCAGACATTTCTGATGTGACATCTGCAATGAGACCGA 720
Db      725 AACCTGATCGTGAGCAGCATCCAGACATTTCTGATGTGACATCTGCAATGAGACCGA 784
Qy      721 GAGGGTTCATGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGCCACCTGGG 780
Db      785 GGGGGTTCATGCGGAGCTACAGCGGACCTTTTCTGAGCCAGGGAGCCACCTGGG 842
Qy      781 ATCTGACCTCTGCGCAACCGGTCACTTTGAGTCCGAGCAGACCCCACTGA 834
Db      843 ATCTGACCTCTGCGCAACCGGTCACTTTGAGTCCGAGCAGACCCCACTGA 896

RESULT 3
LOCUS   BX394295 1201 bp mRNA linear EST 13-MAY-2003
DEFINITION BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC013YB13 5-PRIME, mRNA sequence.
ACCESSION BX394295
VERSION 1
KEYWORDS BX394295.1 GI:30624219
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
L1.W.B., Gruber,C., Jessee,J. and Polyes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC013AC070P1&cluster=4968.r. Contact :
Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope Sequence ID : CS0DC013AC070P1.
FEATURES
source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC013YB13"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was

```

```

BASE COUNT      288 a      284 c      363 g      226 t      40 others
ORIGIN
Query Match      93.4%; Score 779; DB 13; Length 1201;
Best Local Similarity 99.4%; Pred. No. 48-182;
Matches 823; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Qy      1 ATGCTTGGCGGAGGCGAAGATGCGAGACCCCGCGAGGCCGACCGTCCGAC 60
Db      88 ATGCTTGGCGGAGGCGAAGATGCGAGACCCCGCGAGGCCGACCGTCCGAC 146
Qy      61 CAGCGGCTTCTCTGTAAGGGGTGAGCGGCGGCACTCCAGGGGAGAGTGCACCGTGT 120
Db      147 CAGCGGCTTCTCTGTAAGGGGTGAGCGGCGGCACTCCAGGGGAGAGTGCACCGTGT 206
Qy      121 GAGAAATCATGAGATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGGAGAGTGTG 180
Db      207 GAGAAATCATGAGATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGGAGAGTGTG 266
Qy      181 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGCAGACAGAAAGCCAAAGCCTTG 240
Db      267 ATCTGAGCCAGGACAGGTTCTACAAAGTCTGACGCGCAGACAGAAAGCCAAAGCCTTG 326
Qy      241 AAAGCAGTACAAATTTTGCATCCAGTCCAGTGGCTTTGATTAATGATTGATGACAGAGCT 300
Db      327 AAAGCAGTACAAATTTTGCATCCAGTGGCTTTGATTAATGATTGATGACAGAGCT 386
Qy      301 CTGAAGACATGCTGGAGGGCAAAACGGTGAAGTGGCCGACCTATGATTTTGTGACAC 360
Db      387 CTGAAGACATGCTGGAGGGCAAAACGGTGAAGTGGCCGACCTATGATTTTGTGACAC 446
Qy      361 TCAAGTTACAGAGACACAGGTGCTTACCTGCGGACGTGTTCTGTTGAGGGGATC 420
Db      447 TCAAGTTACAGAGACACAGGTGCTTACCTGCGGACGTGTTCTGTTGAGGGGATC 506
Qy      421 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGACAC 480
Db      507 TTGGTGTTCACAGCAGAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGTGACAC 566
Qy      481 GACTCCGACGTGAGCTGTCTCGAAGATTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
Db      567 GACTCCGACGTGAGCTGTCTCGAAGATTCTCCGGGACGTGCGCGGAGGAGGACCTG 626
Qy      541 GAGCAGATTCGACGAGTACACACCTTGTGAAAGCCGCGCTTGCAGAGAGTTTGCCTG 600
Db      627 GAGCAGATTCGACGAGTACACACCTTGTGAAAGCCGCGCTTGCAGAGAGTTTGCCTG 686
Qy      601 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 660
Db      687 CCGACAAAGAGATGCGGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 746
Qy      661 AACCTGATGTCAGACATCCAGACATTTGAAATGAGATCATGCAATGAGGACCGA 720
Db      747 AACCTGATGTCAGACATCCAGACATTTGAAATGAGATCATGCAATGAGGACCGA 806
Qy      721 GAGGGTTCATGCGGAGCTTACAAAGCGACCTTTTCTGAGCCAGGGAGCACCTGGG 780
Db      807 GAGGGTTCATGCGGAGCTTACAAAGCGACCTTTTCTGAGCCAGGGAGCACCTGGG 864
Qy      781 ATGCTGACCTCTGCGCAACCGGTCACTTTGAGTCCAGAGAGACACCC 828
Db      865 ATGCTGACCTCTGCGCAACCGGTCACTTTGAGTCCAGAGAGACACCC 911

RESULT 4
LOCUS   BX343101 1114 bp mRNA linear EST 02-MAY-2003
DEFINITION BX343101 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
CDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
ACCESSION BX343101
VERSION 1
BX343101.1 GI:3034170

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1114)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4968.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL009CB080P1&cluster=4968.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DL009CB080P1.
Location/Qualifiers
1. 1114
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YD15"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_1lb="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (4T) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 247 a 291 c 360 g 202 t 14 others

ORIGIN

Query Match 92.9%; Score 774.4; DB 13; Length 1114;
Best Local Similarity 96.9%; Pred. No. 5.3e-181;
Matches 819; Conservative 1; Mismatches 4; Indels 4; Gaps 4;

QY 1 ATGGCTTCGGGGGAGGAGGAGAGCTGCGAGAGCCCGGCGCGAGCCGACCGTCCGAC 60
Db 159 ATGGCTTCGGGGGAGGAGGAGAGCTGCGAGAGCCCGGCGCGAGCCGACCGTCCGAC 217

QY 61 CAGCGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCGAGGGAAGTGCAGCCGTGTGT 120
Db 218 CAGCGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCGAGGGAAGTGCAGCCGTGTGT 277

QY 121 GAGAGATCATGAGATTGCTGGGACAGAACAGAGTGGAAACGCGGACGCGAAGTGTGTC 180
Db 278 GAGAGATCATGAGATTGCTGGGACAGAACAGAGTGGAAACGCGGACGCGAAGTGTGTC 337

QY 181 ATCTGAGCCGAGAGCAGGTTCTCAAGAGTCTGACGCGAGAGAGGCCAAGGCGCTTG 240
Db 338 ATCTGAGCCGAGAGCAGGTTCTCAAGAGTCTGACGCGAGAGAGGCCAAGGCGCTTG 397

QY 241 AAGAGCAGTCAATTTTGAACCATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 300
Db 398 AAGAGCAGTCAATTTTGAACCATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 457

QY 301 CTGAGAAACATCTGAGAGGGGAAAACGCTGAGAGTGCACCTATGATTTTGTACACAC 360
Db 458 CTGAGAAACATCTGAGAGGGGAAAACGCTGAGAGTGCACCTATGATTTTGTACACAC 517

QY 361 TCAAGTTTACAGAGACCAAGGATGATCCCTGCGAGCGTGTCTGTTGAGGGATC 420
Db 518 TCAAGTTTACAGAGACCAAGGATGATCCCTGCGAGCGTGTCTGTTGAGGGATC 577

QY 421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTTCAACCTGCGCTCTTGTGTGACACC 480
Db 578 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTTCAACCTGCGCTCTTGTGTGACACC 637

QY 481 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGAGAGTGGCCGAGGAGGAGCCTG 540
Db 638 GACTCCGAGCTCAGGCTGTCTCGAAGATTCTCCGGAGAGTGGCCGAGGAGGAGCCTG 697

QY 541 GAGCAGATTCTGACGAGATACACCACTTCGTGAAACCGGCTTCGAGAGATTCTGCTG 600
Db 698 GAGCAGATTCTGACGATTAACACCACTTCGTGAAACCGGCTTCGAGAGATTCTGCTG 757

QY 601 CCGACAAAGAGTATCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 660
Db 758 CCGACAAAGAGTATCCGATGTGATCATCCACGAGAGTGGACATATGTTGCCATC 817

QY 661 AACCTGATGTGAGACATATCCAGACATTTCTGAATGTGACATCTGCAATGGACCGA 720
Db 818 AACCTGATGTGAGACATATCCAGACATTTCTGAATGTGACATCTGCAATGGACCGA 877

QY 721 GGAGGTTCCATAGGGGCGAGCTCAAGCGGACCTTTCTGAGCCAGAGGACCACTGTGG 780
Db 878 GGAGGTTCCATAGGGGCGAGCTCAAGCGGACCTTTCTGAGCCAGAGGACCACTGTGG 935

QY 781 ATCTGACCTCTGGCAACGCTCAATTTGAGTCCAGACAGACCC 828
Db 936 ATCTGACCTCTGGCAACGCTCAATTTGAGTCCAGACAGACCC 982

RESULT 5
B0719741 956 bp mRNA linear EST 16-JUL-2002
AGENCOURT 8229663 lupski dorsal root ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.
B0719741
B0719741.1 GI:21858638
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 956)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13574 row: b column: 18
High quality sequence stop: 552.
Location/Qualifiers
1. 956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev stage="adult, 36 yr"
/lab_host="DH10B"
/clone_1lb="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGGCTCG-3' and 5'-GACTATTTCAATCCGAGCGGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life

FEATURES
source

Db 486 GACTCCGACGTGAGAGTCTCGAAGAGTTCTCCGGGACGTGCGTCGAGGAGGGACCTG 545
 Qy 541 GAGCAGATTCTGAGCGAGTACACACCTTCGTAAGCGCGGCTTCGAGAGTTCTGCTG 600
 Db 546 GAGCAGATTCTGAGCGAGTACACACCTTCGTAAGCGCGGCTTCGAGAGTTCTGCTG 605
 Qy 601 CCGACAAAGAGTATGCGAGTGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
 Db 606 CCGACAAAGAGTATGCGAGTGTATCATCCACAGAGAGTGAACAATATGTTGCCATC 665
 Qy 661 AACCTGATCTGCGACACATCCAGACATTTGTAATGTTGACATCTGCAATGCGACCGA 720
 Db 666 AACCTGATCTGCGACATCATCCAGACATTTGTAATGTTGACATCTGCAATGTTACCGT 725
 Qy 721 GGAGGCTCAATGGCGGAGCTTACAG-CGGACCTTTTCTGAGCGAGGAGCA-CCCTG 778
 Db 726 AGAGGCTCAATGGCGGAGCTTACAGTCCAGACCTTTTCTGAGTCCAGGAGCACACCTG 785
 Qy 779 GGATGCTGACCTCTGCGAAGCG 801
 Db 786 GGATGCTGACCTCTGGAAGCG 808

RESULT 7
 LOCUS BG491358 860 bp mRNA linear EST 27-MAR-2001
 DEFINITION 602536642P1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684728 5',
 mRNA sequence.

ACCESSION BG491358
 VERSION BG491358.1 GI:13452870
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 860)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 DNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
 DNA Sequencing by: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLU at:
 http://image.lnl.gov
 Plate: L1CM1493 row: c column: 11
 High quality sequence stop: 844.

FEATURES
 source 1..860
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684738"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene), Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 195 a 223 c 288 g 154 t
 ORIGIN

Query Match 89.5%; Score 746.8; DB 10; Length 860;
 Best Local Similarity 99.1%; Pred. No. 3.2e-174;
 Matches 803; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 1 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
 Db 52 ATGGCTTCGCGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 111
 Qy 61 CAGCGGCCCTTCTGATAGGGGATGAGCGCGGAGACTGCGCGGAGAGTGCAGTGTGT 120
 Db 112 CAGCGGCCCTTCTGATAGGGGATGAGCGCGGAGACTGCGCGGAGAGTGCAGTGTGT 171
 Qy 121 GAGAGATCATGAGATTGCTGGGACGAAACGAGTGGAAACAGGCGGAGAGGTGTC 180
 Db 172 GAGAGATCATGAGATTGCTGGGACGAAACGAGTGGAAACAGGCGGAGAGGTGTC 221
 Qy 181 ATCTGAGCAGGACGAGTCTCAAGAGTCTTACGCGCAGAGAGAGCCAGCCCTTG 240
 Db 232 ATCTGAGCAGGACGAGTCTCAAGAGTCTTACGCGCAGAGAGAGCCAGCCCTTG 291
 Qy 241 AAAGACAGTACAAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 300
 Db 292 AAAGACAGTACAAATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGACT 351
 Qy 301 CTGAAGAACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 360
 Db 352 CTGAAGAACATCGTGAAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGAACAC 411
 Qy 361 TCAAGTTTACAGAGACCAAGTGTCTTACCTGCGGACGTTGTTGAGGCGATC 420
 Db 412 TCAAGTTTACAGAGACCAAGTGTCTTACCTGCGGACGTTGTTGAGGCGATC 471
 Qy 421 TTGGTGTCTTACAGCCAGAGATCCGAGATGTTCCACTGCGCTCTTGTGACACC 480
 Db 472 TTGGTGTCTTACAGCCAGAGATCCGAGATGTTCCACTGCGCTCTTGTGACACC 531
 Qy 481 GACTCCGACGTGAGCGTGTCTGAG-AGTTCTCCGGAACGTGCGCGAGGAGGACCT 539
 Db 532 GACTCCGACGTGAGCGTGTCTGAGAGAGTTCTCCGGAACGTGCGCGAGGAGGACCT 591
 Qy 540 GGAGCAGATTCTGACGAGTACACACACTTCTGGAAGCGGCTTGAAGAGTTCTGCT 599
 Db 592 GGAGCAGATTCTGACGAGTACACACACTTCTGGAAGCGGCTTGAAGAGTTCTGCT 651
 Qy 600 GCCGACAAAGATATGCGAGTGTATCATCCAGAGAGTGAACATA-TGTTGGCA 658
 Db 652 GCCGACAAAGATATGCGAGTGTATCATCCAGAGAGTGAACATA-TGTTGGCA 711
 Qy 659 TCAACCTGATCGTGAAGACATCCAGAGATTCTGAATGTTGACATCTGAATGGA-C 717
 Db 712 TCAACCTGATCGTGAAGACATCCAGAGATTCTGAATGTTGACATCTGAATGGA-C 771
 Qy 718 CGAGAGAGTTCATGAGGCGGAGCTTCAAGCGGACCTTTTCTGAGCGAGGACACCT 777
 Db 772 CGAGAGAGTTCATGAGGCGGAGCTTCAAGCGGACCTTTTCTGAGCGAGGACACCT 830
 Qy 778 GGGATGCTGACCTCT-GGCAACGCTCACA 806
 Db 831 GGGATGCTGACCTCTGGGCAACGCTCACA 860

RESULT 8
 LOCUS BG826894 916 bp mRNA linear EST 22-MAY-2001
 DEFINITION 602750978P1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
 mRNA sequence.
 ACCESSION BG826894
 VERSION BG826894.1 GI:14174481
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 916)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)


```

Db      52 ATGGCTTGGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGTCGCGAC 111
Qy      61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGCT 120
Db      112 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGCT 171
Qy      121 GAGAAGATCATGAGATTGCTGGGACAGACGAGGTGAAACGCGGACGCGAGAGGTGTC 180
Db      172 GAGAAGATCATGAGATTGCTGGGACAGACGAGGTGAAACGCGGACGCGAGAGGTGTC 231
Qy      181 ATCTGAGCCGAGACAGGTCTCAAGAGTCTGACGCGAGCGAGAGGCCAAGGCTTG 240
Db      232 ATCTGAGCCGAGACAGGTCTCAAGAGTCTGACGCGAGCGAGAGGCCAAGGCTTG 291
Qy      241 AAGAGACGTAATTTTGAACCATCCAGATGCTTGAATTAATGATTGATGACAGAGACT 300
Db      292 AAGAGACGTAATTTTGAACCATCCAGATGCTTGAATTAATGATTGATGACAGAGACT 351
Qy      301 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
Db      352 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 411
Qy      361 TCAAGGTTACCAAGACCAAGGTGTCTACCTCGGACGTGCTTCTGTTGAAGGCAATC 420
Db      412 TCAAGGTTACCAAGACCAAGGTGTCTACCTCGGACGTGCTTCTGTTGAAGGCAATC 471
Qy      421 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 480
Db      472 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 531
Qy      481 GACTCCGACGTCAAGCTGTCTGAAAGTTCTCCGGACGTGCGCCGAGGAGGACCTG 540
Db      532 GACTCCGACGTCAAGCTGTCTGAAAGTTCTCCGGACGTGCGCCGAGGAGGACCTG 591
Qy      541 GAGAGATTTCTGAGCGACATACACACCTTCGTGAAGCCGCGCTTCGAGGAACTTTCCTG 600
Db      592 GAGAGATTTCTGAGCGACATACACACCTTCGTGAAGCCGCGCTTCGAGGAACTTTCCTG 651
Qy      601 CCGACAAAG-AAGTATGCGGATGTGATCATCCAGAGAGTGAACAATAT-GGTTGCA 658
Db      652 CCGACAAAGAGTATGCGGATGTGATCATCCAGAGAGTGAACAATATGCGTTGCA 711
Qy      659 TCAACTGATGTGAGACATCAAGACATTTGAAATGTGACATCTGCAATGGCA-C 717
Db      712 TCAACTGATGTGAGACATCAAGACATTTGAAATGTGACATCTGCAATGGCA-C 771
Qy      718 CGAGAGAGGTCCATTTGGCGGAGCTACAGCGGACCTTTTCTGAGCCGAGGACCACT 777
Db      772 CGAGAGAGGTCCATTTGGCGGAGCTACAGCGGACCTTTTCTGAGCCGAGG-ACCACT 829
Qy      778 GGGATGCTGACCTCTGGCAACGGTCACTTTG 810
Db      830 GGGATGCTGACCTCTGGCAACGGTCACTTTG 862

RESULT 10
LOCUS   AUJ31406 779 bp mRNA linear EST 01-Aug-2002
DEFINITION AUJ31406 NT2RP3 Homo sapiens cDNA clone NT2RP3002519 5', mRNA
ACCESSION AUJ31406
VERSION   AUJ31406.1 GI:10991760
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 779)
Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE     HRI human cDNA project
JOURNAL   Unpublished

```

```

COMMENT Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986
          Email: genomics@hri.co.jp
          HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
          Research Institute; cDNA library construction: Department of
          Virology, Institute of Medical Science, University of Tokyo, and
          Helix Research Institute.
FEATURES
  source
    1..779
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="NT2RP3002519"
    /cell_type="teratocarcinoma"
    /cell_line="NT2"
    /clone_1lb="NT2RP3"
    /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
    cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 179 a 197 c 253 g 147 t 3 others
ORIGIN
Query Match 87.1%; Score 726.8; DB 9; Length 779;
Best Local Similarity 99.1%; Pred. No. 2,7e-169;
Matches 750; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
Qy      1 ATGGCTTGGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGTCCGAC 60
Db      24 ATGGCTTGGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGCCGACCGTCCGAC 83
Qy      61 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGCT 120
Db      84 CAGGGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGCT 143
Qy      121 GAGAAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTGTC 180
Db      144 GAGAAGATCATGAGATTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGTGTGTC 203
Qy      181 ATCTGAGCCGAGACAGGTCTCAAGAGTCTGACGCGAGAGCAAGAGCCGACCTTG 240
Db      204 ATCTGAGCCGAGACAGGTCTCAAGAGTCTGACGCGAGAGCAAGAGCCGACCTTG 263
Qy      241 AAGAGACGTAATTTTGAACCATCCAGATGCTTGAATTAATGATTGATGACAGGACT 300
Db      264 AAGAGACGTAATTTTGAACCATCCAGATGCTTGAATTAATGATTGATGACAGGACT 323
Qy      301 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
Db      324 CTGAAGAACATCGTGAAGGCGAAGAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 383
Qy      361 TCAAGGTTACCAAGACCAAGGTGTCTACCTCGGACGTGCTTCTGTTGAAGGCAATC 420
Db      384 TCAAGGTTACCAAGACCAAGGTGTCTACCTCGGACGTGCTTCTGTTGAAGGCAATC 443
Qy      421 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 480
Db      444 TTGGTGTCTACAGCCAGAGATCCGAGGACATGTTCCACTGCGCTCTTGTGAGACAC 503
Qy      481 GACTCCGACGTCAAGCTGTCTGAAAGTTCTCCGGACGTGCGCCGAGGAGGACCTG 540
Db      504 GACTCCGACGTCAAGCTGTCTGAAAGTTCTCCGGACGTGCGCCGAGGAGGACCTG 563
Qy      541 GAGCAGATTCTGACGAGTACACACCTTGTGTAACCGCGCTTTCGAGAGATTCTGCTG 600
Db      564 GAGCAGATTCTGACGAGTACACACCTTGTGTAACCGCGCTTTCGAGAGATTCTGCTG 623
Qy      601 CCGACAAAGAGATGCCAGATGTGATCATCCACGAGAGGTGGAACAATATGTTGCATC 660
Db      624 CCGACAAAGAGATGCCAGATGTGATCATCCACGAGAGGTGGAACAATATGTTG-CAATC 662

```

Qy 661 AACCTGATCTGAGACATCCAGACATTCGTGATGTGATCATCTGCAATGAGCAGCA 720
 Db 683 AACCTGATCTGCA-CACATTCAGACATTCGTGATGTGATCATCTGCAATGAGCAGCA 741
 Qy 721 GAGGGTCCATGGGCGGAGCTCAAGCGGACCTTTT 757
 Db 742 GGANGGTCCAATGGGCGAGCTCAAGCGGACCTTTT 778

RESULT 11
 BM917506
 LOCUS
 DEFINITION
 AGENCOURT 6605568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
 5', mRNA sequence.
 ACCESSION
 VERSION
 BM917506
 BM917506.1 GI:19367885
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1044)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2012 row: k column: 10
 High quality sequence stop: 624.
 Location/Qualifiers
 1..1044
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5484033"
 /issue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA Synthesis Kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 224 a 306 c 318 g 196 t
 ORIGIN

Query Match 86.9%; Score 724.8; DB 12; Length 1044;
 Best local similarity 98.9%; Pred. No. 9,7e-169;
 Matches 740; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 13 GGAGGGCAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCCCTTC 72
 Db 1 GGAGGGCAAGACTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCCCTTC 60
 Qy 73 CTGATAGGGGTGAGCGCGGCGGACCTGCGAGCGGAAAGTGAACCGTGTGTGAGAAAGATCATG 132
 Db 61 CTGATAGGGGTGAGCGCGGCGGACCTGCGAGCGGAAAGTGAACCGTGTGTGAGAAAGATCATG 120
 Qy 133 GAGTGTCTGGGACAGAAAGAGGTGGAACAGCGGCGGAAAGTGTGATCTGAGCCAG 192
 Db 121 GAGTGTCTGGGACAGAAAGAGGTGGAACAGCGGCGGAAAGTGTGATCTGAGCCAG 180
 Qy 193 GACAGGTTTACAAGGCTCTGACGGCAGACAGAAAGCCAAAGGCTTGAAAGGACAGTAC 252

Db 181 GACAGGTTTACAAGGCTCTGACGGCAGAGCAGAAAGCCAAAGGCTTGAAAGGACAGTAC 240
 Qy 253 AATTGGACATCTCAAGATGCTTTGATATGATTTGATGACAGGACTCTGAAGAACATC 312
 Db 241 AATTGGACATCTCAAGATGCTTTGATATGATTTGATGACAGGACTCTGAAGAACATC 300
 Qy 313 GTGAGAGGCAAAAGCGTGGAGGCGGACCTATGATTTGTGTGACACACTCAAGTTACCA 372
 Db 301 GTGAGAGGCAAAAGCGTGGAGGCGGACCTATGATTTGTGTGACACACTCAAGTTACCA 360
 Qy 373 GAGACCAAGATGATCTTACCGTGGAGCGTGTGTTGTGAGGSCATCTTGTTCTAC 432
 Db 361 GAGACCAAGATGATCTTACCGTGGAGCGTGTGTTGTGAGGSCATCTTGTTCTAC 420
 Qy 433 AGCCAGAGATCCGGGACATGTTTCACTCGGCTCTTGTGTGACACCGACTCCAGCTC 492
 Db 421 AGCCAGAGATCCGGGACATGTTTCACTCGGCTCTTGTGTGACACCGACTCCAGCTC 480
 Qy 493 AGGCTGTCTGGAAGATTCGCGGAGCGTGGCGCGAGGAGGAGACCTGGAGCATTCCTG 552
 Db 481 AGGCTGTCTGGAAGATTCGCGGAGCGTGGCGCGAGGAGGAGACCTGGAGCATTCCTG 540
 Qy 553 AGCCAGTACACACCTTGTGAAAGCCGCTTGAAGAGTTCCTGCGGACCAAGAAAG 612
 Db 541 AGCCAGTACACACCTTGTGAAAGCCGCTTGAAGAGTTCCTGCGGACCAAGAAAG 600
 Qy 613 TATGCCGATGTATCATCCGAGAGGTGAGACATATGTTGCCATCACTGATCTG 672
 Db 601 TATGCCGATGTATCATCCGAGAGGTGAGACATATGTTGCCATCACTGATCTG 660
 Qy 673 CAGCATATCAGGACATTCGTAATGTGTGATCTGCAATGGACAGGAGGTCCTCAT 732
 Db 661 CAGCATATCAGGACATTCGTAATGTGTGATCTGCAATGGACAGGAGGTCCTCAT 720
 Qy 733 GGAG-AGAGCTACAGAGCGGACCTTTTCT 759
 Db 721 GGCGGGAGCTACAGCGGAGCCTTTT 748

RESULT 12
 B0935919
 LOCUS
 DEFINITION
 AGENCOURT 8785353 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6371179
 5', mRNA sequence.
 ACCESSION
 VERSION
 B0935919
 B0935919.1 GI:22351302
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 922)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP/Gazdar
 cDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2544 row: o column: 20
 High quality sequence stop: 625.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6371179"

/tissue type="large cell carcinoma"
 /lab host="DH10B (phage-resistant)"
 /clone lib="NIH MGC 18"
 /note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dt priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 203 a 255 c 297 g 161 t 6 others
 ORIGIN

Query Match 85.6%; Score 714.2; DB 13; Length 922;
 Best Local Similarity 96.1%; Pred. No. 3.9e-165;
 Matches 805; Conservative 0; Mismatches 26; Indels 7; Gaps 7;

QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGACCCCGCGCGGAGCGACCGTCCGAC 60
 Db 36 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGACCCCGCGCGGAGCGACCGTCCGAC 95
 QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTGCAGCTGTGT 120
 Db 96 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTGCAGCTGTGT 155
 QY 121 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACGGCGGACGGGAGGTGTC 180
 Db 156 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACGGCGGAGGTGTC 215
 QY 181 ATCTGAGCCAGGACAGAGTTCTCAAGAGTCTTCAACGAGAGCAAGAGCCAGCTTG 240
 Db 216 ATCTGAGCCAGGACAGAGTTCTCAAGAGTCTTCAACGAGAGCAAGAGCCAGCTTG 275
 QY 241 AAAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTTGATGACAGACT 300
 Db 276 AAAAGACAGTCAATTTTGAACATCCAGATCCCTTTGATATGATTTGATGACAGACT 335
 QY 301 CTGAAGAATCATGAGGAGGCAAAACGCTGAGAGTGGCGGACCTTATGATTTTGTACACAC 360
 Db 336 CTGAAGAATCATGAGGAGGCAAAACGCTGAGAGTGGCGGACCTTATGATTTTGTACACAC 395
 QY 361 TCAAGGTTACAGAGACCAACGAGTGTACCTCGACGAGCTGTCTGTTTGAAGGATC 420
 Db 396 TCAAGGTTACAGAGACCAACGAGTGTACCTCGACGAGCTGTCTGTTTGAAGGATC 455
 QY 421 TTGGTGTCTCAACGCGAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTGACACAC 480
 Db 456 TTGGTGTCTCAACGCGAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTGACACAC 515
 QY 481 GACTCCGACGTCAAGGCTGTCTCAAGAGTCTTCCGGGACGTGGCGGAGGAGGACCTG 540
 Db 516 GACTCCGACGTCAAGGCTGTCTCAAGAGTCTTCCGGGACGTGGCGGAGGAGGACCTG 575
 QY 541 GAGCAGATTCTGACGACGATACCAACCTTCTGTAAGCGCGGCTTCTGAGAGTTCGCTG 600
 Db 576 GAGCAGATTCTGACGACGATACCAACCTTCTGTAAGCGCGGCTTCTGAGAGTTCGCTG 635
 QY 601 CCGACAAAGAGTATGC-CCATGTGATCAT-CCACGAGAGTGG-ACATATAGTGTGCC 657
 Db 636 CCGACAAAGAGTATGCATGTGATCATCCACGAGAGTGGTGMCAATATAGTGTGCC 695
 QY 658 ATCAACCTGATCGAGACATCCAG-GACATTCTGAATGGTGAACATTCGCAATAGGCA 716
 Db 696 ATCAACCTGATCGAGACATCCAGAGACATTCGAAATGTATACATTCGCAATAGGCA 755
 QY 717 CCGA-GGAGGTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCGAGGGAG-CAC 774
 Db 756 CCGAGGAGGTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCGAGGAGCCACAC 815
 QY 775 CTTGGAGTCTGACCTCTGGCAA-ACGGTCACTTTTGGAGTCCAGCAGCAGACCCCAAC 831
 Db 816 CTTGGAGTCTGACCTCTGGCAAACGNNCCCCCATTTTGGAGTCCAGCAAAACCCCCC 873

RESULT 13
 BI258532 848 bp mRNA linear EST 17-JUL-2001
 LOCUS 602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5',
 DEFINITION mRNA sequence.
 ACCESSION BI258532
 VERSION BI258532.1 GI:14814971
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 848)
 NIH-MGC <http://mgc.nci.nih.gov/>,
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHM11271 row: 0 column: 09
 High quality sequence stop: 839.
 Location/Qualifiers

FEATURES

1..848
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5112032"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: 'oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

BASE COUNT 196 a 219 c 271 g 162 t

ORIGIN

Query Match 85.2%; Score 710.2; DB 12; Length 848;
 Best Local Similarity 95.2%; Pred. No. 3.6e-165;
 Matches 797; Conservative 0; Mismatches 8; Indels 32; Gaps 5;

QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGACCCCGCGCGGAGCGACCGTCCGAC 60
 Db 6 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGACCCCGCGCGGAGCGACCGTCCGAC 65
 QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTGCAGCTGTGT 120
 Db 66 CAGCGGCCCTTCCTGATAGG-----GTCGACCGTGTG 98
 QY 121 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACAGGCGGAGGAGGTGTC 180
 Db 99 GAGAAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACAGGCGGAGGAGGTGTC 158
 QY 181 ATCTGAGCCAGGACAGAGTTCTCAAGAGTCTTCAACGAGAGCAAGAGCCAGGCTTG 240
 Db 159 ATCTGAGCCAGGACAGAGTTCTCAAGAGTCTTCAACGAGAGCAAGAGCCAGGCTTG 218
 QY 241 AAAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 300
 Db 219 AAAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGACT 278
 QY 301 CTGAAGAATCATGAGGAGGCAAAACGCTGAGAGTGGCGGACCTTATGATTTTGTGACACAC 360
 Db 279 CTGAAGAATCATGAGGAGGCAAAACGCTGAGAGTGGCGGACCTTATGATTTTGTGACACAC 318

Qy	361	CAAGGTTACAGAGACCAAGCGTGGTCTAACCTCGGGAGCGTGGTCTGTTTGAAGGCATC	420
Db	339	TCAAGGTTACAGAGACCAAGCGTGGTCTAACCTCGGGAGCGTGGTCTGTTTGAAGGCATC	398
Qy	421	TTGGTGTCTACAGGCAGAGAGATCCGGGACATGTTCCACTGGCTCTTCGTGACACC	480
Db	399	TTGGTGTCTACAGGCAGAGAGATCCGGGACATGTTCCACTGGCTCTTCGTGACACC	458
Qy	481	GACTCCGACGTCAGGCTGTCTTCGAAAGATTCTCCGGGACGTGCGCCGAGGAGGGACCTG	540
Db	459	GACTCCGACGTCAGGCTGTCTTCGAAAGATTCTCCGGGACGTGCGCCGAGGAGGGACCTG	518
Qy	541	GAGCAGATTCTGACGCACTACACCACTTCGTGAAGCGGACCTTCGAGGAGTTCTGCTG	600
Db	519	GAGCAGATTCTGACGCACTACACCACTTCGTGAAGCGGACCTTCGAGGAGTTCTGCTG	578
Qy	601	CCGACAAAGAGTATGCCGATGTGTATCATCCACGAGAGTGGACAAATATGTTGCCATC	660
Db	579	CCGACAAAGAGTATGCCGATGTGTATCATCCACGAGAGTGGACAAATATGTTGCCATC	638
Qy	661	AACCTGATCGAGCAGCATCCAGGACATTCGAA--TGGTGCATCTGGAATGGACCG	719
Db	639	AACCTGATCGAGCAGCATCCAGGACATTCGAAATGGTGCATCTGGAATGGACCG	698
Qy	720	AGAGGGTCCAATGGGCGG--AGCTACAGCGGACCTTTCTGAGCCAGGGGACCACTCT	777
Db	699	AGAGGGTCCAATGGGCGGAGCTTACAAGGGGACTTTTCTGAGCCAGGGGACCACTCT	758
Qy	778	GGGATGCTGACCTCTGGCAAC--GGTCACTTTGG--AGTCCAGCAGCAGACCCCACT	832
Db	759	GGGATGCTGACCTCTGGCAACGGGTCACTTTGGAAAGTCCAGCAGCAGACCCCACT	815

RESULT 14					
LOCUS	BO072501				
DEFINITION	BO072501	1036 bp	mRNA	linear	EST 02-APR-2002
	AGENCOURT 6638909	NIH_MGC_122	Homo sapiens	CDNA clone	IMAGE:576169
	5', mRNA sequence.				

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BQ072501.1 GI:19901547
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
 Mammalia; Eulhetaia; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1036)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.

FEATURES

```

1. .1036
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:5761694"
   /lab_host="DH10B"
   /clone_1b="NH MGc 122"
   /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6
   Site_1: NciI; Site 2: EcoRV (destroyed); RNA source
   anonymous pool of 72 week female lung, 16 week female
   spleen, and 20-22 week male spleens. Library is oligo-dT
   primed and directionally cloned (EcoRV site is destroyed

```

upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC library."

Query Match	85.0%	Score 709;	DB 13;	Length 1036;
Best Local Similarity	97.2%	Pred. No. 7.9e15;		
Matches 765; Conservative	0;	Mismatches 15;	Indels 7;	Gaps 4;

QY	1	ATGGCTTCGGCGGAGGCGAAGA	CTGCAGAGGCCCGCGCGGAGGCGACGCTCCGAC	60
Db	64	ATGGCTTCGGCGGAGGCGAAGA	CTGCAGAGGCCCGCGCGGAGGCGACGCTCCGAC	123
QY				
QY	61	CAGCGGCCCTTCTCTGATAGGG	GTGAGCGGCGGCACTGCCAGCGGGAAAGTCGACCTGTGT	120
Db	124	CAGCGGCCCTTCTCTGATAGGG	GTGAGCGGCGGCACTGCCAGCGGGAAAGTCGACCTGTGT	183
QY				
QY	121	GAGAAGATCATGGAGTTGCTGG	GAACGAAGAGTGGAAACAGGGCGGCGGGAAGTGGTTC	180
Db	184	GAGAAGATCATGGAGTTGCTGG	GAACGAAGAGTGGAAACAGGGCGGCGGGAAGTGGTTC	243
QY	181	ATCCTGAGCCAGACAGAGTTCT	CAAGAGTCTCTACCGCAGAGCAGAAAGCCAGGCCCTTG	240
Db	244	ATCCTGAGCCAGACAGAGTTCT	CAAGAGTCTCTACCGCAGAGCAGAAAGCCAGGCCCTTG	303
QY	241	AAAGGACAGTACAATTTTGAC	ATCCATCCAGATGCTCTTGATATGATTTGATGACAGAGACT	300
Db	304	AAAGGACAGTACAATTTTGAC	ATCCATCCAGATGCTCTTGATATGATTTGATGACAGAGACT	363
QY	301	CTGAAAGACATCGTGGAGGGC	CAAAACGCTGGAGGTGCGGACCTATGATTTTGTGACACAC	360
Db	364	CTGAAAGACATCGTGGAGGGC	CAAAACGCTGGAGGTGCGGACCTATGATTTTGTGACACAC	423
QY	361	TCAGAGTTACACAGAGACCA	CGGTGTCTACCTCGCGACGTGGTTCGTTTGAAGGAGATC	420
Db	424	TCAGAGTTACACAGAGACCA	CGGTGTGTACCTCGCGACGTGGTTCGTTTGAAGGAGATC	483
QY	421	TTGGTGTCTACAGCCAGAG	AGATCCGGGACATGTTCCACTCGGCTCTTCTGTGACAC	480
Db	484	TTGGTGTCTACAGCCAGAG	AGATCCGGGACATGTTCCACTCGGCTCTTCTGTGACAC	543
QY	481	GACTCCGAGTCAAGGCTGTCT	GGAAGAATTCCTCGGAGCGTGGCGCGAAGGAGGACCTG	540
Db	544	GACTCCGAGTCAAGGCTGTCT	GGAAGAATTCCTCGGAGCGTGGCGCGAAGGAGGAGCCTG	603
QY	541	GAGCAGATTCTGACGCGAGTA	CAACACCTTCGTGAAGCGGCTTCGAGAGTTCTGCCTG	600
Db	604	GAGCAGATTCTGACGCGAGTA	CAACACCTTCGTGAAGCGGCTTCGAGAGTTCTGCCTG	663
QY	601	-CCGACAAAGAGTATGCCGAT	GTGATCATCCACAGAGAGTGGACAAATATGTTGCCAT	659
Db	664	-CCGACAAAGAGTATGCCGAT	GTGATCATCCACAGAGAGTGGACAAATATGTTGCCAT	723
QY	660	CAACCTGATCTGGACAGACAT	-CCAGAGCAATTCGAAATGGTGCATCTGCAAAATGGGACC	718
Db	724	CAACCTGATCTGGACAGACAT	-CCAGAGCAATTCGAAATGGTGCATCTGCAAAATGGGACC	783
QY	719	GAGG--AGGGTCCAAATGGG	CGAGACTACAAGCGGACCTTTTCTGAG--CCAGGGGACCA	773
Db	784	GAGGAGGGGTCAACATGGG	GGAAGCTACAAAGCGGACCTTTTCTGAGCCAGGGGACCA	843
QY	774	CCCTGGG	780	
Db	844	CCCTGG	850	

RESULT 15	BI261258	LOCUS	827 bp	mRNA	linear	EST 17-Jul-2001
DEFINITION	602869110F1 NIH MGC_12	Homo sapiens cDNA clone IMAGE:5108666 5',				

mRNA sequence.
 ACCESSION B1261258
 VERSION B1261258.1 GI:14820349
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgsabp@remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/MLNL at: http://image.lnl.gov
 Plate: LAM1263 Row: C Column: 03
 High quality sequence start: 5
 High quality sequence stop: 799.
 Location/Qualifiers
 1..827
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5108666"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 12"
 /note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 Kb. Library prepared by Life Technologies."
 BASE COUNT 189 a 210 c 274 g 154 t
 ORIGIN
 Query Match 84.9%; Score 708.2; DB 12; Length 827;
 Best Local Similarity 97.2%; Pred. No. 1.1e-164;
 Matches 753; Conservative 0; Mismatches 18; Indels 4; Gaps 3;
 QY 1 ATGGCTTCGGGCGGAGGCGAAGCTGCGAGAGCCCGCGGAGCGGACCGTCCGCAC 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 52 ATGGCTTCGGGCGGAGGCGAAGCTGCGAGAGCCCGCGGAGCGGACCGTCCGCAC 111
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTCGACCGTGT 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 112 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTCGACCGTGT 171
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 GAGAAATCATGAGTCTGCGGACAGACGAGGTGAAACAGCGGACGCGAGGTGTC 180
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 172 GAGAAATCATGAGTCTGCGGACAGACGAGGTGAAACAGCGGACGCGAGGTGTC 231
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 ATCTGAGCCGAGCAGGTTCTACAAGTCTGACGCGAGGAGGAGGCGGAGGCTTG 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 232 ATCTGAGCCGAGCAGGTTCTACAAGTCTGACGCGAGGAGGAGGCGGAGGCTTG 291
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 AAAGGACAGTAAATTTTGAATCCATCCAGATGCCCTTGTATATATGATGACAGGACT 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 292 AAAGGACAGTAAATTTTGAATCCATCCAGATGCCCTTGTATATATGATGACAGGACT 351
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 CTGAAGAACATCTGAGAGGCGAAACGCTGAGGTGCCGACTATGATTTTGTGACAC 360
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 352 CTGAAGAACATCTGAGAGGCGAAACGCTGAGGTGCCGACTATGATTTTGTGACAC 411
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 TCAAGTTACAGAGACCAAGGTGCTACCCCTGCGGACGCGTTCGTTTGAAGGCATC 420
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 412 TCAAGTTACAGAGACCAAGGTGCTACCCCTGCGGACGCGTTCGTTTGAAGGCATC 471
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTGACACACC 480
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 472 TTGGTGTCTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTGACACACC 531
 QY 481 GACTCCGACGTCAAGGTCTGTTCGAAAGATTCT-CCGGAGCTGCGCGAGGAGGACT 539
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 532 GACTCCGACGTCAAGGTCTGTTCGAAAGATTCTCCCGGAGCTGCGCGAGGAGGACT 591
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 540 GAGGACATTCGACGAGTACCAACCTTCGAAAGCGGCGCTTGAGAGGTTGCGCT 599
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 592 GAGGACATTCGACGAGTACCAACCTTCGAAAGCGGCGCTTGAGAGGTTGCGCT 651
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 600 GCCGACAAAGATATGCGGATGTGATCCACAGAGAGTGAACATATGTTGCGCAT 659
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 652 GCCGACAAAGATATGCGGATGTGATCCACAGAGAGTGAACATATGTTGCGCAT 711
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 660 CAACCTGATGTGACGACAT--CCAGGACATTTGAT-GTGACATCTGCAATGCGA 716
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 712 CAACCTGATGTGACGACATTTCCAGGACATTTGATGAGTGGTGAATCTGCAATGCGA 771
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 717 CCGAGGAGGTGCAATGAGGCGGAGCTCAAGCGGACCTTTCTGAGCGAGGGAC 771
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 772 CCGAGGAGGTGCAATGAGGCGGAGCTCAAGCGGACCTTTCTGAGCGAGGGAC 826
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: November 25, 2003, 02:00:51
 Job time : 2012.96 secs

Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MSC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHC82012 row: k column: 10
High quality sequence stop: 624

FEATURES
source

Location/Qualifiers
1..1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5484033"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_106"
/note="Organ: blood; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 224 a 306 c 318 g 196 t

ORIGIN

Query Match 84.9%; Score 708; DB 12; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

13 GGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGACCAAGCGCCCTTC 72
1 GGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGACCAAGCGCCCTTC 60
73 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGAGAAAGATCATG 132
61 CTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGAGAAAGATCATG 120
133 GAGTTCCTGGGACGAACGAGGTGGAACGCGGCAAGGTGTGATCTTGAAGCAG 192
121 GAGTTCCTGGGACGAACGAGGTGGAACGCGGCAAGGTGTGATCTTGAAGCAG 180
193 GACAGGTTCTACAGAGTCTGAGCGGAGAGAGGCGGCTTGAAGAGCAGTAC 252
181 GACAGGTTCTACAGAGTCTGAGCGGAGAGAGGCGGCTTGAAGAGCAGTAC 240
253 AATTTGACCATCCAGATGCTTGTATATATTTGATGCAAGAGCTCGAAGACATC 312
241 AATTTGACCATCCAGATGCTTGTATATATTTGATGCAAGAGCTCGAAGACATC 300
313 GTGAGGCGCAAAAGGTGAGAGTCCGACCTATATTTTGTGACACACTCAAGTTTACA 372
301 GTGAGGCGCAAAAGGTGAGAGTCCGACCTATATTTTGTGACACACTCAAGTTTACA 360
373 GAGACCAAGGTGCTACCTCGGAGAGTGTCTGTTTGAAGGACATCTTGTTTAC 432
361 GAGACCAAGGTGCTACCTCGGAGAGTGTCTGTTTGAAGGACATCTTGTTTAC 420
433 AGCAGAGAGATCCGAGCATGTTTCACCTGCGCTCTTGTGACACCGACTCCGAGTCC 492
421 AGCAGAGAGATCCGAGCATGTTTCACCTGCGCTCTTGTGACACCGACTCCGAGTCC 480
493 AGGCTGTCTGGAAGAGTTCTCCGAGAGTGTGCGGAGAGAGGACTGTGACAGATTCTG 552
481 AGGCTGTCTGGAAGAGTTCTCCGAGAGTGTGCGGAGAGAGGACTGTGACAGATTCTG 540
553 AGCAGAGATCCGAGCATGTTTCACCTGCGGCTTGTGAGAGGATCTGCGCGGACAAAGAG 612
541 AGCAGAGATCCGAGCATGTTTCACCTGCGGCTTGTGAGAGGATCTGCGCGGACAAAGAG 600
613 TATGCCGATGATCATCCGAGAGAGTGTGCAATATGTTTGCATCAACCTGATGTTG 672
601 TATGCCGATGATCATCCGAGAGAGTGTGCAATATGTTTGCATCAACCTGATGTTG 660
673 CAGGACATCCGAGCATTTCTGAATGTGACATCTTGCAGAAATGGACCGA 720
661 CAGGACATCCGAGCATTTCTGAATGTGACATCTTGCAGAAATGGACCGA 708

RESULT 2
BM467984

LOCUS 1118 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532885
5', mRNA sequence.
ACCESSION BM467984
VERSION BM467984.1 GI:18517026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1118)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
UNPUBLISHED
AUTHORS Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgabs@mail.nih.gov
COMMENT Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
Plate: LLM12216 row: n column: 22
High quality sequence stop: 689.

FEATURES
source

Location/Qualifiers
1..1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/issue_type="telomerasecoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

BASE COUNT 234 a 311 c 361 g 209 t 3 others

ORIGIN

Query Match 82.7%; Score 690; DB 12; Length 1118;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCTTGGCGGAGCGGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 60
47 ATGAGCTTGGCGGAGCGGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGAC 106
61 CAGCGGCGCTCTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGT 120
107 CAGCGGCGCTCTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACCGTGTGT 166
121 GAGAGATCATGAGTGTGCTGGGAGAGAGAGTGTGGAACAAGCGGAGAGAGTGTGTC 180
167 GAGAGATCATGAGTGTGCTGGGAGAGAGAGTGTGGAACAAGCGGAGAGAGTGTGTC 226
181 ATCTTAGGCGAGGACAGGTTCTCAAGGTCCTGACGCGAGAGGCAAGGCGCTTG 240
227 ATCTTAGGCGAGGACAGGTTCTCAAGGTCCTGACGCGAGAGGCAAGGCGCTTG 286
241 AAAGGACATGATTTTGAACCATCCAGATGCTTGTATATATGATTTGATGACAGAGACT 300
287 AAAGGACATGATTTTGAACCATCCAGATGCTTGTATATATGATTTGATGACAGAGACT 346
301 CTGAGAGATCATGAGGAGGCAAAAGGTGAGAGTGTGCGGACCTATATATTTTGTGACACAC 360
347 TCAAGGTTACAGAGACCAAGGTGTCTACCGTCCGAGAGTGTGTTGTTGAGGCGATC 420
407 TCAAGGTTACAGAGACCAAGGTGTCTACCGTCCGAGAGTGTGTTGTTGAGGCGATC 466

QY	421	TTGGTGTCTCA	CAGCAGAGAAATCCGGACAATGTTCCACCTGGCGCTCTTGTTGGACAC	480
Db	467	TTGGTGTCTCTCA <td>CAGCCAGGAAGATCCGGACATGTTTCACTGGCGCTCTTGTTGGACAC</td> <td>526</td>	CAGCCAGGAAGATCCGGACATGTTTCACTGGCGCTCTTGTTGGACAC	526
QY	481	GACTCCGACGT <td>CAAGGCTGTCTCGAAGATTTCTCCGGAGCGTGCCGCGAGGGAGGACCTG</td> <td>540</td>	CAAGGCTGTCTCGAAGATTTCTCCGGAGCGTGCCGCGAGGGAGGACCTG	540
Db	527	GATTCGACG <td>CGTCAAGGCTGTCTCGAAGATTTCTCCGGAGCGTGCGCGAGGGAGGAGCCTG</td> <td>586</td>	CGTCAAGGCTGTCTCGAAGATTTCTCCGGAGCGTGCGCGAGGGAGGAGCCTG	586
QY	541	GAGCAGATTTCTGA <td>CGCAGTACCACTTCTGTGAACCGCGCTTGAGGAGTTCTGCTG</td> <td>600</td>	CGCAGTACCACTTCTGTGAACCGCGCTTGAGGAGTTCTGCTG	600
Db	587	GAGCAGATTTCTGA <td>CGCAGTACCACTTCTGTGAACCGCGCTTGAGGAGTTCTGCTG</td> <td>646</td>	CGCAGTACCACTTCTGTGAACCGCGCTTGAGGAGTTCTGCTG	646
QY	601	CCGACAAAGAA <td>GTATGCGCATGTGTCATCCACAGAGAGTGAGCAATATGTTGGCCATC</td> <td>660</td>	GTATGCGCATGTGTCATCCACAGAGAGTGAGCAATATGTTGGCCATC	660
Db	647	CCGACAAAGAA <td>GTATGCGCATGTGTCATCCACAGAGAGTGAGCAATATGTTGGCCATC</td> <td>706</td>	GTATGCGCATGTGTCATCCACAGAGAGTGAGCAATATGTTGGCCATC	706
QY	661	AACCTGATCGT <td>GCAGCAGCATCCAGGACATT</td> <td>690</td>	GCAGCAGCATCCAGGACATT	690
Db	707	AACCTGATCGT <td>GCAGCAGCATCCAGGACATT</td> <td>736</td>	GCAGCAGCATCCAGGACATT	736
RESULT 3				
CD514811		898 bp	mRNA	linear
LOCUS				
DEFINITION				
AGNCOURT	14376196	NIH MGC 181	Homo sapiens	CDNA clone
IMAGE	30356201	5', mRNA	sequence.	
CD514811				
VERSION	CD514811.1	GI:31446529		
KEYWORDS	EST.			
SOURCE				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
1 (bases 1 to 898)				
NIH-MGC	http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished				
Contact: Daniela S. Gerhard, Ph.D.				
Office of Cancer Genomics				
National Cancer Institute / NIH				
Bldg. 31 Rm10A07 Bethesda, MD 20892				
Email: GSapbs-r@mail.nih.gov				
Tissue Procurement: Dr. Michael Brownstein				
CDNA Library Preparation: Invitrogen Corp				
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)				
DNA Sequencing by: Agencourt Bioscience Corporation				
Clone distribution: MGC clone distribution information can be				
found through the I.M.A.G.E. Consortium/LNLN at:				
http://image.llnl.gov				
Plate: NDAM477	row: 1	column: 10		
High quality sequence stop: 662.				
Location/Qualifiers				
1. .898				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="IMAGE:30396201"				
/tissue_type="White Matter"				
/dev_stage="Unknown"				
/lab_host="ID10B-Ton A (T1 and T5 phage resistances) "				
/clone_lib="NIH MGC 181"				
/note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV				
(destroyed); Library is oligo-dT primed and directionally				
cloned (EcoRV site is destroyed upon cloning) . Average				
insert size 1.42 kb. Library was constructed by				
(Invitrogen) . Note: this is a NIH MGC Library."				
BASE COUNT	194 a	231 c	310 g	158 t
ORIGIN				
Query Match	82.0%	Score 684;	DB 14;	Length 898;
Best Local Similarity	100.0%;	Pos. NO. 0;		
Matches 684;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGGCTTCGGGGGGAGGCGAAGACTGGGAGAGCCCGCGCGGAGGCGACACCGTCCGAC	60
Db	72	ATGGCTTCGGGGGGAGGCGAAGACTGGGAGAGCCCGCGCGGAGGCGACACCGTCCGAC	131
QY	61	CAGCGGCCCTTCCCTGATATAGGGGTGAGCGGCGGCGACCTGCGAGCGGGAAAGTCAGCCGTGTGT	120
Db	132	CAGCGGCCCTTCCCTGATATAGGGGTGAGCGGCGGCGACCTGCGAGCGGGAAAGTCAGCCGTGTGT	191
QY	121	GAGAAAGTCAATGAGATTGCTGGGACAGAACGAGGTGAAACAGCGGACAGCGGAAGTGTGT	180
Db	192	GAGAAAGTCAATGAGATTGCTGGGACAGAACGAGGTGAAACAGCGGACAGCGGAAGTGTGT	251
QY	181	ATCTGGAGCCAGGACAGGTTCTAACAAGTCTCTGACGCGAGAGCGAAGGCCAAGGCTCTTG	240
Db	252	ATCTGGAGCCAGGACAGGTTCTAACAAGTCTCTGACGCGAGAGCGAAGGCCAAGGCTCTTG	311
QY	241	AAAGGACAGTACATTTTGAACCATCCAGATGCCCTTTATATGATTTGATGACAGGACT	300
Db	312	AAAGGACAGTACATTTTGAACCATCCAGATGCCCTTTATATGATTTGATGACAGGACT	371
QY	301	CTGAAGAACATCGTGAAGGGCAAAACGTTGAGGTGCGGACCTATGATTTTGTGACACAC	360
Db	372	CTGAAGAACATCGTGAAGGGCAAAACGTTGAGGTGCGGACCTATGATTTTGTGACACAC	431
QY	361	TCAAGGTTACAGAGACCAAGGTGTCTAACCTGCGGACGTTGATTTTGAAGGACATC	420
Db	432	TCAAGGTTACAGAGACCAAGGTGTCTAACCTGCGGACGTTGATTTTGAAGGACATC	491
QY	421	TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCCTTCTGTGAGACAC	480
Db	492	TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCCCTTCTGTGAGACAC	551
QY	481	GACTCCGACGTGAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCCGAGGGAGGAGACTG	540
Db	552	GACTCCGACGTGAGGCTGTCTCGAAGATTCTCCGGGACGTGCGCCCGAGGGAGGAGACTG	611
QY	541	GAGCAGATTCTGACGCGAGTACACCACTTGTGTGAAGCGGCGCTTGAGGAATTGTGCGCTG	600
Db	612	GAGCAGATTCTGACGCGAGTACACCACTTGTGTGAAGCGGCGCTTGAGGAATTGTGCGCTG	671
QY	601	CCGACAAAGAGATATGCCGATGTGATCATCCAGAGAGTGAACAAATATGTTGGCATC	660
Db	672	CCGACAAAGAGATATGCCGATGTGATCATCCAGAGAGTGAACAAATATGTTGGCATC	731
QY	661	AACCTGATCGTGACGACATCCGAG	684
Db	732	AACCTGATCGTGACGACATCCGAG	755
RESULT 4			
LOCUS	BI256928	725 bp	mRNA linear EST 17-JUL-2001
DEFINITION	60297553.F01 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114816 5',		
ACCESSION	BI256928		
VERSION	BI256928.1	GI:14811813	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: Incyte Genomics, Inc.		
	CDNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		

found through the I.M.A.G.E. Consortium/LINL at:

http://image.lnl.gov
Plate: LHM11279 row: c column: 09
High quality sequence stop: 707.

FEATURES

source

1..725
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5114816"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH MGC 12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by life
Technologies."

BASE COUNT 175 a 181 c 224 g 145 t
ORIGIN

Query Match 81.2%; Score 677; DB 12; Length 725;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 GGAAGTGAACCGTGTGAGAGATCATGAGTGTGGGACAGAGAGGTGGAAGC 163
DB 1 GGAAGTGAACCGTGTGAGAGATCATGAGTGTGGGACAGAGAGGTGGAAGC 60
QY 164 GGCAGCGGAAGGTGATCTCTGAGCAGAGCAGGTTCTACAGAGTCTTGCAGCAGC 223
DB 61 GGCAGCGGAAGGTGATCTCTGAGCAGAGCAGGTTCTACAGAGTCTTGCAGCAGC 120
QY 224 AGAAGCCCAAGGCTTGAAGGACAGTCAATTTTGAACATCCAGATGCTTTGATATG 283
DB 121 AGAAGCCCAAGGCTTGAAGGACAGTCAATTTTGAACATCCAGATGCTTTGATATG 180
QY 284 ATTGATGACACAGGACTCTGAAGAACATCGTGAAGGCAAAAGGTGAGGTGCGACT 343
DB 181 ATTGATGACACAGGACTCTGAAGAACATCGTGAAGGCAAAAGGTGAGGTGCGACT 240
QY 344 ATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTCTACCTGCGGACGTG 403
DB 241 ATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTCTACCTGCGGACGTG 300
QY 404 TTCTGTTTGAAGGATCTTGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGC 463
DB 301 TTCTGTTTGAAGGATCTTGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGC 360
QY 464 GCTCTCTCGTGAACCGGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTG 523
DB 361 GCTCTCTCGTGAACCGGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTG 420
QY 524 GCCAGGAGGAGCCTGAGACGATTTCTGACGAGTACACCACTTGTGAGCCGCGCT 583
DB 421 GCCAGGAGGAGCCTGAGACGATTTCTGACGAGTACACCACTTGTGAGCCGCGCT 480
QY 584 TCGAGAGGTTTGTCTGCTGCGGACAAAGAGATGCGCATGTGATATCCACAGAGAGTGG 643
DB 481 TCGAGAGGTTTGTCTGCTGCGGACAAAGAGATGCGCATGTGATATCCACAGAGAGTGG 540
QY 644 ACAATATGTTGATCATCACTGATCGTGAAGCAGCATCCAGAGACATTCGTAATGTTGACA 703
DB 541 ACAATATGTTGATCATCACTGATCGTGAAGCAGCATTCGTAATGTTGACA 600
QY 704 TCTGCAATGACACGAGAGAGGTCAATGAGCGGAGCTACAGCGGACCTTTTCTGAGC 763
DB 601 TCTGCAATGACACGAGAGAGGTCAATGAGCGGAGCTACAGCGGACCTTTTCTGAGC 660
QY 764 CAGGGAGCACCTCTGGG 780
DB 661 CAGGGAGCACCTCTGGG 677

RESULT 5

BG826894

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

916 bp mRNA linear EST 22-MAY-2001
602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
mRNA sequence.
BG826894
BG826894.1 GI:14174481
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 916)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.lnl.gov
Plate: LHM1802 row: k column: 20
High quality sequence stop: 843.

FEATURES

source

1..916
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4903819"
/tissue_type="rhodomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 17"
/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 203 a 239 c 309 g 165 t
ORIGIN

Query Match 77.9%; Score 650; DB 12; Length 916;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGGAGACCTGCGAGAGCCCCCGCGGAGCCGACCGTCCGAC 60
DB 80 ATGGCTTCGGCGGAGGCGGAGACCTGCGAGAGCCCCCGCGGAGCCGACCGTCCGAC 139
QY 61 CAGCGGCCCTCTCTGATAGGGGTGAGCGCGGACCTCGAGCGGAGAGTGCACCGGTCT 120
DB 140 CAGCGGCCCTCTCTGATAGGGGTGAGCGCGGACCTCGAGCGGAGAGTGCACCGGTCT 199
QY 121 GAGAGATCATGAGGTTGCTGGGACGAAAGAGGTGGAACAGCGGCGAGCGGAAGTGTGTC 180
DB 200 GAGAGATCATGAGGTTGCTGGGACGAAAGAGGTGGAACAGCGGCGAGCGGAAGTGTGTC 259
QY 181 ATCTGAGCGGAGACAGGTTCTTCAAGAGTCTTGAACGCGAGACGAGACGAGGCGCAAGCCTTTG 240
DB 260 ATCTGAGCGGAGACAGGTTCTTCAAGAGTCTTGAACGCGAGACGAGGCGCAAGCCTTTG 319
QY 241 AAGGACAGTACAAATTTTGAACATCCAGATGCTTTGATGATTTGATTTGATGACAGAGACT 300
DB 320 AAGGACAGTACAAATTTTGAACATCCAGATGCTTTGATGATTTGATTTGATGACAGAGACT 379
QY 301 CTGAAGAACATCGTGAAGGCAAAAGGTGAGGTGCCAGCTATGATTTTGTGACACAC 360

Db 380 CTGAGAACATCGTGGAGGCGCAAAACGGTGGAGGTGCCGACTATGATTTTGTGACAC 439
Qy 361 TCAAGTTACCAAGACCAAGGTTGTTACCTTCGCGACGTGGTTCTGTTTGAAGGATC 420
Db 440 TCAAGGTTACCAAGACCAAGGTTGTTACCTTCGCGACGTGGTTCTGTTTGAAGGATC 499
Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTTCCAGCTGGCCCTTTCGTGACAC 480
Db 500 TTGGTGTTCACAGCCAGAGATCCGGACATGTTTCCAGCTGGCCCTTTCGTGACAC 559
Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGAGACGTGCGCCAGAGGAGGACCTG 540
Db 560 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGAGACGTGCGCCAGAGGAGGACCTG 619
Qy 541 GAGCAGATTCTGACGACGTACACCACTTGTGTAAGCCGCTTCGAGAGTTCTGCTG 600
Db 620 GAGCAGATTCTGACGACGTACACCACTTGTGTAAGCCGCTTCGAGAGTTCTGCTG 679
Qy 601 CCGACAAAGAGTATGCGATGTGATCCACAGAGAGTGACAATAT 650
Db 680 CCGACAAAGAGTATGCGATGTGATCCACAGAGAGTGACAATAT 729

RESULT 6
BX394295 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION BX394295
ACCESSION BX394295
VERSION BX394295.1 GI:30624219
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
14, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC013AC07QPLcluster=4968.r. Contact :
Peng liang Email : filiang@lifetech.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC013AC07QPL.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC013E13"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-clio(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 288 a 284 c 363 g 226 t 40 others
ORIGIN

Query Match 77.1% Score 643; DB 13; Length 1201;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 46 GCCGACCGTCCGACCAAGCGCCCTTCTGATAGGGGTGAGCGCGCACTGCGACGGG 105
Db 132 GCCGACCGTCCGACCAAGCGCCCTTCTGATAGGGGTGAGCGCGCACTGCGACGGG 191

Qy 106 AAGTCGACCGTGTGAGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCG 165
Db 192 AAGTCGACCGTGTGAGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCG 251
Qy 166 CAGCGAAGGTGTGATCTCTGAGCCAGACAGATTCTCAAGGTCTCTGACGCGAGAC 225
Db 252 CAGCGAAGGTGTGATCTCTGAGCCAGACAGATTCTCAAGGTCTCTGAGCGAGAC 311
Qy 226 AAGCCGACGCTTGAAGAGACATCAATTTTGAACATCCAGATCCCTTGAATGAT 285
Db 312 AAGCCGACGCTTGAAGAGACATCAATTTTGAACATCCAGATCCCTTGAATGAT 371
Qy 286 TTGATGACACGACCTTGAAAGAACATCGTGAGAGGAGAAACGTTGAGGCGGACCTAT 345
Db 372 TTGATGACACGACCTTGAAAGAACATCGTGAGAGGAGAAACGTTGAGGCGGACCTAT 421
Qy 346 GATTTTGTGACACATCAAGTTACAGAGACCAAGGTGTCTACCTCGGACGTGTT 405
Db 432 GATTTTGTGACACATCAAGTTACAGAGACCAAGGTGTCTACCTCGGACGTGTT 491
Qy 406 CTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGATCCGAGACATGTTCCACTGCGC 465
Db 492 CTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGATCCGAGACATGTTCCACTGCGC 551
Qy 466 CTCTTGTGACACCGACCTCCGACCTGACGCTGTCTCCGAAAGTTCTCCGAGCGTGGCG 525
Db 552 CTCTTGTGACACCGACCTCCGACCTGACGCTGTCTCCGAAAGTTCTCCGAGCGTGGCG 611
Qy 526 CGAGGAGGAGACCTGAGAGATTTGACGACGATACCAACCTTCGTGAAGCCGCGCTTC 585
Db 612 CGAGGAGGAGACCTGAGAGATTTGACGACGATTTGACGACGATACCAACCTTCGTGAAGCCGCGCTTC 671
Qy 586 GAGGAGTTCTGCTGCTGCGCAAAAGATATGCCATGTGATATCCACAGAGAGTGAC 645
Db 672 GAGGAGTTCTGCTGCTGCGCAAAAGATATGCCATGTGATATCCACAGAGAGTGAC 721
Qy 646 AATATGTTGCCATCAACTGATCGTGACGACATCTCAGACATTTGAAATGTTGACATC 705
Db 732 AATATGTTGCCATCAACTGATCGTGACGACATCTCAGACATTTGAAATGTTGACATC 791
Qy 706 TGCATATGACACCGAGAGGTTCCATATGGGCGGA 739
Db 792 TGCATATGACACCGAGAGGTTCCATATGGGCGGA 825

RESULT 7
BG770518 730 bp mRNA linear EST 15-MAY-2001
LOCUS BG770518
DEFINITION 602734226F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4859577 5',
mRNA sequence.
ACCESSION BG770518 GI:14081171
VERSION BG770518
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 730)
1 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC/DC/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (JLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/JLNL at:
http://image.jnl.gov
Plate: LLCML715 row: h column: 10
High quality sequence stop: 720.

FEATURES
source

Location/Qualifiers
1. .730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4859577"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 49"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 156 a 190 c 258 g 126 t

Query Match 76.7%; Score 640; DB 12; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
DB 82 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 141
QY 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTCCGACGGGAAAGTGCAGCTGTGT 120
DB 142 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTCCGACGGGAAAGTGCAGCTGTGT 201
QY 121 GAGAAGATCATGAGTTCCTGGACAGAACGAGGTGAAACGCGGACGCGGAGAGTGTGTC 180
DB 202 GAGAAGATCATGAGTTCCTGGACAGAACGAGGTGAAACGCGGACGCGGAGAGTGTGTC 261
QY 181 ATCTGAGCCGAGGACAGGTTCTCAAGGTCTCTGACGCGAGCAGAAAGCCGACCTTG 240
DB 262 ATCTGAGCCGAGGACAGGTTCTCAAGGTCTCTGACGCGAGCAGAAAGCCGACCTTG 321
QY 241 AAAGAGACATGATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 300
DB 322 AAAGAGACATGATTTTGAACATCCAGATCCCTTTGATTAATGATTTGATGACAGACT 381
QY 301 CTGAAGAATCGTGGAGGCGCAAAACGCTGGAGGCGGACCTATGATTTTGTGACAC 360
DB 382 CTGAAGAATCGTGGAGGCGCAAAACGCTGGAGGCGGACCTATGATTTTGTGACAC 441
QY 361 TCAAGGTTACCGAGACCCAGGTGTCTACCTCGGACGTCGATTTCTTTGAGGGATC 420
DB 442 TCAAGGTTACCGAGACCCAGGTGTCTACCTCGGACGTCGATTTCTTTGAGGGATC 501
QY 421 TTGGTCTTCAACACGAGGAGATCCGGGACATGTTTCCACCTCGGCTCTTCTGTGACAC 480
DB 502 TTGGTCTTCAACACGAGGAGATCCGGGACATGTTTCCACCTCGGCTCTTCTGTGACAC 561
QY 481 GACTCCAGCTCAGGCTGTCTCAAGAGTTCTCCGGACGTCGCGGAGGAGGAGGACCTG 540
DB 562 GACTCCAGCTCAGGCTGTCTCAAGAGTTCTCCGGACGTCGCGGAGGAGGAGGACCTG 621
QY 541 GAGCAGATTCTGACGACGATACCACTTCGTGAAGCCGAGCTTCGAGAGATTCTGCTG 600
DB 622 GAGCAGATTCTGACGACGATACCACTTCGTGAAGCCGAGCTTCGAGAGATTCTGCTG 681
QY 601 CCGACAAAGAGTATGCCGATGATCATCCCAAGAGAG 640
DB 682 CCGACAAAGAGTATGCCGATGATCATCCCAAGAGAG 721

RESULT 8
BO719741
LOCUS BO719741 956 bp mRNA linear EST 16-JUL-2002

DEFINITION
AGENCOURT 8229663 lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.
BO719741
VERSION BO719741.1 GI:21858638
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Plate: LHMJ3574 row: b column: 18
High quality sequence stop: 552.

FEATURES
source

Location/Qualifiers
1. .956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_1ib="lupski_dorsal_root_ganglion"
/note="Vector: pCMV-SPORT (Life Technologies); Site 1: NotI; Site 2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCCG-3' and 5'-GACTAGTTCAGATCGGAGCGGCGGCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 202 a 272 c 296 g 186 t

Query Match 74.2%; Score 619; DB 13; Length 956;
Best Local Similarity 100.0%; Pred. No. 2,3e-309;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 GAGGCGAAGCTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCC 73
DB 20 GAGGCGAAGCTGCGAGAGCCCGCGCGGAGGCGGACCGTCCGACCGAGCGGCGCTTCC 79
QY 74 TGATAGGGGTGAGCGGCGGACCTGCGAGCGGAGTGCACCGTGTGTGAGAAATCATGG 133
DB 80 TGATAGGGGTGAGCGGCGGACCTGCGAGCGGAGTGCACCGTGTGTGAGAAATCATGG 139
QY 134 AGTTGCTGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGATCTTGAGCGAG 193
DB 140 AGTTGCTGGACAGAACGAGGTGAAACAGCGGACGCGAAGGTGATCTTGAGCGAG 199
QY 194 ACAGGTTCTTCAAGGTTCTTGAACGCGGACGACAGAAAGCCCAAGCCTTGAAGAACA 253
DB 200 ACAGGTTCTTCAAGGTTCTTGAACGCGGACGACAGAAAGCCCAAGCCTTGAAGAACA 259
QY 254 ATTTGACCATCAGATGCTTTGATTAATGATTTGATGACAGAGCTCTGAAGACATCG 313
DB 260 ATTTGACCATCAGATGCTTTGATTAATGATTTGATGACAGAGCTCTGAAGACATCG 319
QY 314 TGAGGCGCAAAACGATGAGGTGCGGACCTATGATTTGTGACACACTCAAGGTTACAG 373

Db 320 TGGAGGGCAAAACGGTGGAGGTGCGACCTATGATTTTGTGACACCTCAAGGTTACGAG 379

Qy 374 AGACCAAGGTGCTTACCTCGGACCTGCTGTTTGAGGGATCTTGTTGTTTACA 433

Db 380 AGACCAAGGTGCTTACCTCGGACCTGCTGTTTGAGGGATCTTGTTGTTTACA 439

Qy 434 GCCAGAGATCCGGGACATGTTCCACCTGCGCCTCTTGAGACACGACCTCGACGTCA 493

Db 440 GCCAGAGATCCGGGACATGTTCCACCTGCGCCTCTTGAGACACGACCTCGACGTCA 499

Qy 494 GGTCTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGAGACCTTGAGCAGATTCTGA 553

Db 500 GGTCTCTGAAAGATTCTCCGGGACGTGCGCCGAGGAGGAGACCTTGAGCAGATTCTGA 559

Qy 554 CGCAGTACACCACTTGTGTGAAGCCGCTTCGAGGAGTTCTGCTCTCCGACAAAGAGT 613

Db 560 CGCAGTACACCACTTGTGTGAAGCCGCTTCGAGGAGTTCTGCTCTCCGACAAAGAGT 619

Qy 614 ATGCCGATGATCATCC 632

Db 620 ATGCCGATGATCATCC 638

RESULT 9
BG390519 989 bp mRNA linear EST 12-MAR-2001
LOCUS 60241618.F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524308 5',
DEFINITION mRNA sequence.
ACCESSION BG390519
VERSION BG390519.1 GI:13283967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 989)
NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM0428 row: f column: 21
High quality sequence stop: 750.
Location/Qualifiers
1. 989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4524308"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NCI;
Site_2: Sali; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 256 a 258 c 314 g 161 t
ORIGIN

Query Match 73.7%; Score 615; DB 10; Length 989;
Best Local Similarity 100.0%; Pired. No. 2.8e-307;
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 CGAAGACTGAGAGCCCGCGCGGAGCGGACCTGCGACAGCGGCCCTTCGTGAT 77

Db 1 CGAAGACTGAGAGCCCGCGCGGAGCGGACCTGCGACAGCGGCCCTTCGTGAT 60

Qy 78 AGGGGTGAGCGGGGACATCGCCAGCGGAAAGTGAACCGTGTGAGAAATCATGAGTT 137

Db 61 AGGGGTGAGCGGGGACATCGCCAGCGGAAAGTGAACCGTGTGAGAAATCATGAGTT 120

Qy 138 GCTGGGACAGAAAGAGTGAACAGCGGACCGGAAAGTGTGATCTTGAGCCAGACAG 197

Db 121 GCTGGGACAGAAAGAGTGAACAGCGGACCGGAAAGTGTGATCTTGAGCCAGACAG 180

Qy 198 GTTCTCAAGGTCCTGACCGCAGAGAGAGGCAAGGCTTGAAGAGACATGACATTT 257

Db 181 GTTCTCAAGGTCCTGACCGCAGAGAGAGGCAAGGCTTGAAGAGACATGACATTT 240

Qy 258 TGACCATCCAGATGCTTTGATATATATTGATGACACAGACTCTGAAGAACTGTTGA 317

Db 241 TGACCATCCAGATGCTTTGATATATATTGATGACACAGACTCTGAAGAACTGTTGA 300

Qy 318 GGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGAC 377

Db 301 GGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACACTCAAGGTTACAGAGAC 360

Qy 378 CACGTTGCTTACCTGCGGACCTGCTGTTTGAAGGATCTTGATTTTACAGCCA 437

Db 361 CACGTTGCTTACCTGCGGACCTGCTGTTTGAAGGATCTTGATTTTACAGCCA 420

Qy 438 GGAATTCGGGACATGTTCCACTGCGCTCTTGAGACACGACCTCGAGCTCAGGCT 497

Db 421 GGAATTCGGGACATGTTCCACTGCGCTCTTGAGACACGACCTCGAGCTCAGGCT 480

Qy 498 GTCTCAAGGTCCTGCGGACCTGCGGAGGAGGAGGACCTTGAGCAGATTCTGACGCA 557

Db 481 GTCTCAAGGTCCTGCGGACCTGCGGAGGAGGAGGACCTTGAGCAGATTCTGACGCA 540

Qy 558 GTACACCACTTCTGTAAGCGGCTTCGAGAGATTCTGCTGCGCAGCAAAAGATATGC 617

Db 541 GTACACCACTTCTGTAAGCGGCTTCGAGAGATTCTGCTGCGCAGCAAAAGATATGC 600

Qy 618 CGATGATCATCC 632

Db 601 CGATGATCATCC 615

RESULT 10
BG491384 869 bp mRNA linear EST 27-MAR-2001
LOCUS 602535670.F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
DEFINITION mRNA sequence.
ACCESSION BG491384
VERSION BG491384.1 GI:13452896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 869)
NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHCMI493 row: k column: 17
High quality sequence stop: 863.
Location/Qualifiers
1. 869
source

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4684936"
/lissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH MGC_41"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."
BASE COUNT      197 a      223 c      291 g      156 t      2 others
ORIGIN
Query Match      73.3%; Score 611; DB 10; Length 869;
Best Local Similarity 100.0%; Pred. No. 3.2e-305;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
Db      52 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 111
QY      61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCGGAAGTTCGACCGTGTGT 120
Db      112 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCGGAAGTTCGACCGTGTGT 171
QY      121 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGGAGGTGTC 180
Db      172 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGGAGGTGTC 231
QY      181 ATCTGAGCCGAGCAGGCTTCAAGAGTCTGACGCGAGAGCGAAGGCCAAGCCCTTG 240
Db      232 ATCTGAGCCGAGCAGGCTTCAAGAGTCTGACGCGAGAGCGAAGGCCAAGCCCTTG 291
QY      241 AAAGAGACGTAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
Db      292 AAAGAGACGTAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 351
QY      301 CTGAAGAACATGTGAGAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTACAC 360
Db      352 CTGAAGAACATGTGAGAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTACAC 411
QY      361 TCAAGGTTACAGAGACCAACGCTGTCTACACCTCGAGAGCTGTTCTTTGAGGCAATC 420
Db      412 TCAAGGTTACAGAGACCAACGCTGTCTACACCTCGAGAGCTGTTCTTTGAGGCAATC 471
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
Db      472 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 531
QY      481 GACTCCGACGTGAGGCTGTCTGAAAGTTTCCGGGACGTGCGCCGAGAGGAGGACTG 540
Db      532 GACTCCGACGTGAGGCTGTCTGAAAGTTTCCGGGACGTGCGCCGAGAGGAGGACTG 591
QY      541 GAGAGATTTCAGAGCGAGTACACCACTTCGTGAAGCCGGGCTTCGAGAGATTCTGCTG 600
Db      592 GAGAGATTTCAGAGCGAGTACACCACTTCGTGAAGCCGGGCTTCGAGAGATTCTGCTG 651
QY      601 CCGACAAAGAA 611
Db      652 CCGACAAAGAA 662

```

```

RESULT 11
AUI31406      779 bp      mRNA      linear      EST 01-AUG-2002
LOCUS      AUI31406 NT2RP3 Homo sapiens cDNA clone NT2RP3002519 5', mRNA
DEFINITION      sequence.
ACCESSION      AUI31406

```

```

VERSION      AUI31406.1 GI:10991760
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 779)
Oca,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE      HRI human cDNA project
JOURNAL      Unpublished
COMMENT      Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel.: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3002519"
/cell_type="leukocarcinoma"
/clone_1b="NT2"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      179 a      197 c      253 g      147 t      3 others
ORIGIN
Query Match      72.5%; Score 605; DB 9; Length 779;
Best Local Similarity 99.8%; Pred. No. 3.9e-302;
Matches 655; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 60
Db      24 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCCGACCGTCCGAC 83
QY      61 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCGGAAGTTCGACCGTGTGT 120
Db      84 CAGCGGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCGGAAGTTCGACCGTGTGT 143
QY      121 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGGAGAGTGTGTC 180
Db      144 GAGAGATCATGAGTTCTGGGACAGAACGAGGTGAAACGCGGACGGAGAGTGTGTC 203
QY      181 ATCTGAGCCGAGCAGGTTCTACAGAGTCTTACCGCAGAGACGAAGGCCAAGCCCTTG 240
Db      204 ATCTGAGCCGAGCAGGTTCTACAGAGTCTTACCGCAGAGACGAAGGCCAAGCCCTTG 263
QY      241 AAAGAGACGTAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 300
Db      264 AAAGAGACGTAATTTTGAACATCCAGATGCTTTGATATGATTTGATGACAGAGCT 323
QY      301 CTGAAGAACATGTGAGAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTACACAC 360
Db      324 CTGAAGAACATGTGAGAGGCAAAACGGTGAAGGTGCGGACCTATGATTTTGTACACAC 383
QY      361 TCAAGGTTACAGAGACCAACGCTGTCTACACCTCGAGAGCTGTTCTTTGAGGCAATC 420
Db      384 TCAAGGTTACAGAGACCAACGCTGTCTACACCTCGAGAGCTGTTCTTTGAGGCAATC 443
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
Db      444 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTCTGTGACACC 503

```

QY 481 GACTCCGACGTACGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACTG 540
| | | | |
Db 504 GACTCCGACGTACGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACTG 563
| | | | |
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTAAAGCCGCTTCGAGAGTTCTGCTG 600
| | | | |
Db 564 GAGCAGATTCTGACGAGTACACCACTTCTGTAAAGCCGCTTCGAGAGTTCTGCTG 623
| | | | |
QY 601 CCGACAAAGAGATGCGGATGTATATCCACGAGAGTGAACATATGTTGC 656
| | | | |
Db 624 CCGACAAAGAGATGCGGATGTATATCCACGAGAGTGAACATATGTTGC 679
| | | | |

RESULT 12
LOCUS B0072501 1036 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694
5', mRNA sequence.
ACCESSION B0072501
VERSION B0072501.1 GI:19901547
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1036)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12810 row: 1 column: 15
High quality sequence stop: 644.
Location/Qualifiers
1. 1036

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5761694"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
site_1: Nott; Site 2: EcorV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcorV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC library."
BASE COUNT 221 a 285 c 346 g 184 t
ORIGIN

Query Match 72.2%; Score 602; DB 13; Length 1036;
Best Local Similarity 100.0%; Pred. No. 166-300;
Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGCGGAGCGAAGACTGCGAGAGCCCGCGCGAGCGCGACCTCCGAC 60
| | | | |
Db 64 ATGGCTTCGCGGAGCGAAGACTGCGAGAGCCCGCGCGAGCGCGACCTCCGAC 123
| | | | |
QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCTGTGT 120
| | | | |
Db 124 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTGCACCTGTGT 183
| | | | |

QY 121 GAGAAATCATGAGATTGCTGGGACAGAAAGAGTGAACAGCGGACGGAAGTGTGTC 180
| | | | |
Db 184 GAGAAATCATGAGATTGCTGGGACAGAAAGAGTGAACAGCGGACGGAAGTGTGTC 243
| | | | |
QY 181 ATCTGAGCCAGAGACAGGTTCTTCAAGAGTCTTGACGCGAGAGCAAGGCCAAGCCCTTG 240
| | | | |
Db 244 ATCTGAGCCAGAGACAGGTTCTTCAAGAGTCTTGACGCGAGAGCAAGGCCAAGCCCTTG 303
| | | | |
QY 241 AAAGACATATCAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGCAAGACT 300
| | | | |
Db 304 AAAGACATATCAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGCAAGACT 363
| | | | |
QY 301 CTGAGAACATCATGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGAACAC 360
| | | | |
Db 364 CTGAGAACATCATGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGAACAC 423
| | | | |
QY 361 TCAAGTTACAGAGACACAGTGTCTTCACTGCGGACGTGTGTTGTTGAGGCAATC 420
| | | | |
Db 424 TCAAGTTACAGAGACACAGTGTCTTCACTGCGGACGTGTGTTGTTGAGGCAATC 483
| | | | |
QY 421 TTGCTGTTTCAACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCTGACAC 480
| | | | |
Db 484 TTGCTGTTTCAACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCTGACAC 543
| | | | |
QY 481 GACTCCGACGTACGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACTG 540
| | | | |
Db 544 GACTCCGACGTACGAGTGTCTCCAGAGATTCTCCGGACCTGCGCGAGGAGGACTG 603
| | | | |
QY 541 GAGCAGATTCTGACGAGTACACCACTTCTGTAAAGCCGCTTCGAGAGTTCTGCTG 600
| | | | |
Db 604 GAGCAGATTCTGACGAGTACACCACTTCTGTAAAGCCGCTTCGAGAGTTCTGCTG 663
| | | | |
QY 601 CC 602
| |
Db 664 CC 665
| |

RESULT 13
LOCUS B1258532 848 bp mRNA linear EST 17-JUL-2001
DEFINITION 602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5',
mRNA sequence.
ACCESSION B1258532
VERSION B1258532.1 GI:14814971
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 848)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11271 row: 0 column: 09
High quality sequence stop: 839.
Location/Qualifiers
1. 848

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5112032"
/issue_type="Cervical carcinoma cell line"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_12"

/note="Organ: cervix; Vector: PCMV-SPORT6; Site:1; Nct1;
Site_2: Sali; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT

196 a 219 c 271 g 162 t

Query Match 70.6%; Score 589; DB 12; Length 848;
Best Local Similarity 100.0%; Pred. No. 7.9e-294;
Matches 589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GTCGACCGTGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGCA 167
DB 86 GTCGACCGTGTGTGAGAGATCATGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGCA 145
QY 168 GCGGAAGTGTCTCTCTGAGCCGAGACAGGTCTTACAGGCTCTGACGCGACGCGAA 227
DB 146 GCGGAAGTGTCTCTCTGAGCCGAGACAGGTCTTACAGGCTCTGACGCGACGCGAA 205
QY 228 GCGCAAGGCTTGAAGAGACAGTCAATTTTGACCATCCAGATGCTTTGATTAATGATTT 287
DB 206 GCGCAAGGCTTGAAGAGACAGTCAATTTTGACCATCCAGATGCTTTGATTAATGATTT 265
QY 288 GATGACAGGACTCTGAGAACATCTGAGAGGCAAAACGTTGAGGCTGCGACCTATGA 347
DB 266 GATGACAGGACTCTGAGAACATCTGAGAGGCAAAACGTTGAGGCTGCGACCTATGA 325
QY 348 TTTTGTGACACACTCAAGGTTTACAGAGACCAAGGTGTCTACCTTGGGACGTTGTTCT 407
DB 326 TTTTGTGACACACTCAAGGTTTACAGAGACCAAGGTGTCTACCTTGGGACGTTGTTCT 385
QY 408 GTTTGAGGAGATCTGTGTTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGGCGCT 467
DB 386 GTTTGAGGAGATCTGTGTTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGGCGCT 445
QY 468 CTTCGTGAGACACGACTCCGACGTCAGAGCTGTCTGAGAGATTCTCCGGACGTTGCCG 527
DB 446 CTTCGTGAGACACGACTCCGACGTCAGAGCTGTCTGAGAGATTCTCCGGACGTTGCCG 505
QY 528 AGGAGAGGACTCTGAGACAGATTTGAGCGCACTACCACTTCTGTAAGCCGCTTGA 587
DB 506 AGGAGAGGACTCTGAGACAGATTTGAGCGCACTACCACTTCTGTAAGCCGCTTGA 565
QY 588 GGAATTCGCTGCGCAAAAGAGTATGCGATGTATCATCCACAGAGAGTGA 647
DB 566 GGAATTCGCTGCGCAAAAGAGTATGCGATGTATCATCCACAGAGAGTGA 625
QY 648 TATGTTGACATCAACTGATCTGAGACATCCAGAGACATTTGTAAT 696
DB 626 TATGTTGACATCAACTGATCTGAGACATCCAGAGACATTTGTAAT 674

RESULT 14

BM545603

LOCUS 1055 bp mRNA linear EST 20-FEB-2002

DEFINITION

AGENCOURT 6500279 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587864
5' mRNA sequence.

ACCESSION

BM545603.1 GI:18777841

VERSION

EST.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

1 (bases 1 to 1055)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL

Unpublished
Contact: Robert Strauberg, Ph.D.
Email: cgabos-f@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L14M12358 row: a column: 17
High quality sequence stop: 579.

FEATURES

source

1..1055
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5587864"
/lab_host="DH10B"
/clone_lib="NIH MGC 125"
/note="Organ: ovary (pool of 3); Vector: PCMV-SPORT6;
Site_1: EcorV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dt primed and directionally cloned
(EcorV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

BASE COUNT

208 a 343 c 314 g 190 t

Query Match 69.8%; Score 582; DB 12; Length 1055;
Best Local Similarity 100.0%; Pred. No. 3.6e-290;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTCGCGGAGGAGCGAAGACTGAGAGCCCGCGCGGAGCGACCGCTCCGAC 60
DB 49 ATGCTTCGCGGAGGAGCGAAGACTGAGAGCCCGCGCGGAGCGACCGCTCCGAC 108
QY 61 CAGCGGCCCTTCTCTGATAGGAGTGAAGCGGCGGCACTGCGAGCGGAAGTGAACGTTGTT 120
DB 109 CAGCGGCCCTTCTCTGATAGGAGTGAAGCGGCGGCACTGCGAGCGGAAGTGAACGTTGTT 168
QY 121 GAGAAGTATGAGAGTTGCTGGGACAGAACAGAGTGAACAGCGGACGCGGAAGTGTGTC 180
DB 169 GAGAAGTATGAGAGTTGCTGGGACAGAACAGAGTGAACAGCGGACGCGGAAGTGTGTC 228
QY 181 ATCTGAGCGAGGACAGGTTCTCAAGGTCCTGACGCGAGAGCAAGGCGGCTTG 240
DB 229 ATCTGAGCGAGGACAGGTTCTCAAGGTCCTGACGCGAGAGCAAGGCGGCTTG 288
QY 241 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATTAATGATTGAAGCAGACT 300
DB 289 AAAGGACAGTACATTTTGAACATCCAGATGCTTTGATTAATGATTGAAGCAGACT 348
QY 301 CTGAAAGACATGCTGAGAGGCAAAACGTTGAGAGTCCGACCTATGATTTTGTGACAC 360
DB 349 CTGAAAGACATGCTGAGAGGCAAAACGTTGAGAGTCCGACCTATGATTTTGTGACAC 408
QY 361 TCAAGTTACAGAGACACAGGTTGCTCAAGCTCCGCGAGCGTGGTCTGTTTGAAGGCACT 420
DB 409 TCAAGTTACAGAGACACAGGTTGCTCAAGCTCCGCGAGCGTGGTCTGTTTGAAGGCACT 468
QY 421 TTGCTGTTTCAAGCGAGAGATCCGGAATGCTTCACTGAGCTCTTCTGTGACAC 480
DB 469 TTGCTGTTTCAAGCGAGAGATCCGGAATGCTTCACTGAGCTCTTCTGTGACAC 528
QY 481 GACTCCGAGTCAAGGTTGCTGAGAGAGTTTCCGGAAGTCCGCGAGAGGAGGACCTG 540
DB 529 GACTCCGAGTCAAGGTTGCTGAGAGAGTTTCCGGAAGTCCGCGAGAGGAGGACCTG 588
QY 541 GAGCAGATTCTGAGCGAGTACACCACTTCTGTAAGCGGCGCC 582
DB 589 GAGCAGATTCTGAGCGAGTACACCACTTCTGTAAGCGGCGCC 630

RESULT 15

BI830878

LOCUS BI830878 661 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603081002F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5172432 5',
 mRNA sequence.
 ACCESSION BI830878
 VERSION BI830878.1 GI:15942428
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 661)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM11429 row: d column: 01
 High quality sequence stop: 661.
 Location/Qualifiers
 1..661
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5172432"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_119"
 /note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH MGC Library."
 BASE COUNT 138 a 173 c 738 g 112 t
 ORIGIN
 Query Match 69.5%; Score 580; DB 12; Length 661;
 Best Local Similarity 100.0%; Pred. No. 3.4e-289;
 Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTTCGGCGGAGGCGAAGCTGGAGAGCCCGCGCGGAGCGCGCGCGCGCAC 60
 DB 82 ATGGCTTCGGCGGAGGCGAAGCTGGAGAGCCCGCGCGGAGCGCGCGCGCGCAC 141
 QY 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACGTGTG 120
 DB 142 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGCACGTGTG 201
 QY 121 GAGAGATCATGAGGTGCTGGAGCAGAACGAGGTGGAACAGCGGCAAGCGGAGGTGCTC 180
 DB 202 GAGAGATCATGAGGTGCTGGAGCAGAACGAGGTGGAACAGCGGCAAGCGGAGGTGCTC 261
 QY 181 ATCTTGAGCCAGACAGGTTCTACAGGTCTTGAACGAGCAGAGCAAGGCGCAAGGCTTGG 240
 DB 262 ATCTTGAGCCAGACAGGTTCTACAGGTCTTGAACGAGCAGAGCAAGGCGCAAGGCTTGG 321
 QY 241 AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGGACT 300
 DB 322 AAAGCAGATCAATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGGACT 381
 QY 301 CTGAAGAACATCGTGAAGGGCAAAACGCTGAGAGTGCAGACTATGATTTTGTGACACAC 360
 DB 382 CTGAAGAACATCGTGAAGGGCAAAACGCTGAGAGTGCAGACTATGATTTTGTGACACAC 441

QY 361 TCAAGTTACAGAGACCAAGGTTGCTACCCGCGGAGCTGTTCTGTTGAGGCGATC 420
 DB 442 TCAAGTTACAGAGACCAAGGTTGCTACCCGCGGAGCTGTTCTGTTGAGGCGATC 501
 QY 421 TTGTTGTTACAGCAGAGAGATCCGAGCATGTTCCACTCGCTCTTGTGACACC 480
 DB 502 TTGTTGTTACAGCAGAGAGATCCGAGCATGTTCCACTCGCTCTTGTGACACC 561
 QY 481 GACTCCGAGCTGAGCTGTTCTGAGAGATTCTCCGAGAGCTGCGCGAGGAGGAGCTTG 540
 DB 562 GACTCCGAGCTGAGCTGTTCTGAGAGATTCTCCGAGAGCTGCGCGAGGAGGAGCTTG 621
 QY 541 GAGCAGATTCTGACGAGTACACCACTTCGTGAAGCCGG 580
 DB 622 GAGCAGATTCTGACGAGTACACCACTTCGTGAAGCCGG 661

Search completed: November 25, 2003, 06:51:16
 Job time : 2012.46 secs

This Page Blank (uspto)


```
Db 141 LeuValPheTyrSerGlnGluIleArgAspMetHeHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCGACGTCAGCGTGTCTCGAGA 507
Db 161 AsperaspValArgLeuSerArgArg 169

RESULT 2
Q9QY68 PRELIMINARY; PRT; 105 AA.
ID Q9QY68
AC Q9QY68;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Uridine kinase (fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=Spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yaoi T., Matsube S., Okajima S., Hirasawa Y., Fushiki S.;
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BAA83085.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
SQ SEQUENCE 105 AA; 11967 MW; 90B43C15E07EB029 CRC64;

Alignment Scores:
Pred. No.: 1.23e-16 Length: 105
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-3 (1-834) x Q9QY68 (1-105)
QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAAGAAGTATGCCGATGTG 624
Db 36 ThrPheValIysProIaIaPheGlnGluPheCysLeuProThrIlySlyrAlaAspVal 55
QY 625 ATCATCCACAGAGA 639
Db 56 IleIleProArgIly 60

RESULT 3
Q9BU42 PRELIMINARY; PRT; 111 AA.
ID Q9BU42
AC Q9BU42;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to uridine monophosphate kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strauberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

Db 141 LeuValPheTyrSerGlnGluIleArgAspMetHeHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCGACGTCAGCGTGTCTCGAGA 507
Db 161 AsperaspValArgLeuSerArgArg 169

RESULT 4
Q8C476 PRELIMINARY; PRT; 261 AA.
ID Q8C476
AC Q8C476;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Uridine-cytidine kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK082837; BAC38646.1; -.
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:
Pred. No.: 1.14e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-3 (1-834) x Q8C476 (1-261)
QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAAGAAGTATGCCGATGTG 624
Db 187 ThrPheValIysProIaIaPheGlnGluPheCysLeuProThrIlySlyrAlaAspVal 206
QY 625 ATCATCCACAGAGA 639
Db 207 IleIleProArgIly 211

RESULT 5
Q92528 PRELIMINARY; PRT; 111 AA.
ID Q92528
AC Q92528;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
```

DT 01-FBB-1997 (TREMBlrel. 02, last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 DB 5'-terminal region of UMK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=96411689; PubMed=8812458;
 RA Ozaki K., Kuroki T., Hayashi S., Nakamura Y.;
 RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
 RT a differential mRNA display method.";
 RL Genomics 36:316-319 (1996).
 DR EMBL; D78335; BA11349.1; -;
 DR InterPro; IPR006083; PRK_URK.
 DR Pfam; PF00485; PRK_1.
 DR PRINTS; PR00988; URIDINKINASE.
 SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Alignment Scores:
 Pred. No.: 3.41e-06 Length: 111
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.42% Indels: 0
 DB: 4 Gaps: 0

US-09-896-522-3 (1-834) x Q92528 (1-111)

QY 565 ACCTGCTGAGCGGCGCTCGAGAGTTCGCTGCGGACAAAG 609
 DB 37 ThrpheVallyspRoAlapheGluGluPhcysleuProThrlys 51

RESULT 6
 ID Q8MRJ1 PRELIMINARY; PRT; 260 AA.
 AC Q8MRJ1;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
 DE ID13909p.
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fiske E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanevong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AY119583; AAWS0237.1; -;
 DR FlyBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK_1.
 DR PRINTS; PR00988; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16E5 CRC64;

Alignment Scores:
 Pred. No.: 3.18e-06 Length: 260
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.42% Indels: 0
 DB: 5 Gaps: 0

US-09-896-522-3 (1-834) x Q8MRJ1 (1-260)

QY 388 TACCCTGCGAGCGTGTGTTTGAGGGCATCTGTGTTCTAC 432
 DB 135 TyrProAlaAspValValleuPhcGluGlyIleLeuValPheTyr 149

Search completed: November 25, 2003, 07:58:09
 Job time: 40.5285 secs

This Page Blank (uspto)

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)
10529.410 Million cell updates/sec

Title: US-09-896-522-3

Sequence: 1 atggtctcggcgagcgca.....ccagcagcagaccctaactga 834

Scoring table: OLIGO_NUC

Searched: 2888711 seqs, 20454813386 residues

word size : 15

Total number of hits satisfying chosen parameters: 29369

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq Length: 200000000000

Post-processing: Listing first 45 summaries

Database :

```

1:  gb_ba.*
2:  gb_hcg.*
3:  gb_in.*
4:  gb_om.*
5:  gb_ov.*
6:  gb_pac.*
7:  gb_ph.*
8:  gb_pl.*
9:  gb_pr.*
10: gb_rc.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_in.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_to.*
27: em_scs.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rad.*
36: em_hcg_mod.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	834	100.0	834	6	AX449219	AX449219 Sequence
2	834	100.0	1624	6	AX449217	AX449217 Sequence
3	834	100.0	2160	6	BD157613	BD157613 Primer fo
4	834	100.0	2160	6	AK022317	AK022317 Homo sapi
5	783	93.9	834	9	AF254133	AF254133 Homo sapi
6	783	93.9	1022	9	AF237290	AF237290 Homo sapi
7	575	68.9	753	9	BD146824	BD146824 Primer fo
8	517	62.0	1395	9	AF125106	AF125106 Homo sapi
9	508	60.9	2072	9	BC015547	BC015547 Homo sapi
10	497	59.6	2228	9	AK057848	AK057848 Homo sapi
11	338	40.5	734	6	AX540411	AX540411 Sequence
12	183	21.9	147492	9	AL358781	AL358781 Human DNA
13	36	4.3	660	11	BD077668	BD077668 S212P6227
14	36	4.3	1810	10	MUSURK1	U31783 Mus musculu
15	36	4.3	1959	10	BC025146	BC025146 Mus muscu
16	36	4.3	233210	2	AC097693	AC097693 Rat muscu
17	36	4.3	2411882	10	AL0808027	AL0808027 Mouse DNA
18	36	3.6	244105	2	AC098897	AC098897 Rattus no
19	36	3.6	254677	2	AC068494	AC068494 Mus muscu
20	33	4.0	254953	9	AC078885	AC078885 Mus muscu
21	33	2.8	336	9	BT006860	BT006860 Homo sapi
22	23	2.8	336	12	BT007548	BT007548 Synthetic
23	23	2.8	383	11	GA1891	GA1891 SHGC-35183
24	23	2.8	744	9	AB062451	AB062451 Homo sapi
25	23	2.8	1121	9	AF236637	AF236637 Homo sapi
26	23	2.8	1209	9	BC002906	BC002906 Homo sapi
27	23	2.8	1322	6	AX135546	AX135546 Sequence
28	23	2.8	99866	9	AL358115	AL358115 Human DNA
29	23	2.8	164246	2	AC016371	AC016371 Homo sapi
30	22	2.6	233137	2	AC127758	AC127758 Rattus no
31	21	2.5	841	11	BD017303	BD017303 S212P6359
32	21	2.5	952	10	MUSURK1A	U31784 Mus musculu
33	21	2.5	3060	6	OC010358	U10358 Oryctolagus
34	21	2.5	3060	6	I13759	I13759 Sequence 1
35	21	2.5	192336	2	AC118474	AC118474 Mus muscu
36	21	2.5	211075	10	AC076974	AC076974 Mus muscu
37	20	2.4	655	11	BD077340	BD077340 S212P6009
38	20	2.4	1581	5	AF004318	AF004318 Carassius
39	20	2.4	1784	5	BC045968	BC045968 Dario rer
40	20	2.4	4663	1	AY181035	AY181035 Methyloba
41	20	2.4	62299	2	AC100224	AC100224 Mus muscu
42	20	2.4	133036	2	AC109801_3	Continuation (4 of
43	20	2.4	148865	2	AC101821	AC101821 Mus muscu
44	20	2.4	148865	2	AL954171	AL954171 Dario rer
45	20	2.4	162978	2	AC132083	AC132083 Mus muscu

ALIGNMENTS

RESULT 1				
AX449219	AX449219	834 bp	DNA	linear PAT 03-JUL-2002
LOCUS	Sequence 3 from Patent WO0202761.			
DEFINITION	AX449219			
ACCESSION	AX449219.1	GI:21697996		
VERSION				
KEYWORDS	.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Glucksmann,M.A.			
TITLE	57658, a human uridine kinase and uses thereof			
JOURNAL	Patent: WO 0202761-A 3 10-JUN-2002;			

Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 834;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGAGAACTCGAGAGCCCCCGCGGAGGCGCAACCTCCGAC 60
DB 1 ATGGCTTCGGGGGAGGAGAACTCGAGAGCCCCCGCGGAGGCGCAACCTCCGAC 60
QY 61 CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGT 120
DB 61 CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGT 120
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
DB 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
QY 181 ATCTGAGCCGAGACAGGTTCTCAAGAGTCTGACGCGAGACGAGAGCCAAAGCCCTTG 240
DB 181 ATCTGAGCCGAGACAGGTTCTCAAGAGTCTGACGCGAGACGAGAGCCAAAGCCCTTG 240
QY 241 AAAGGACATCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGAGCT 300
DB 241 AAAGGACATCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGAGCT 300
QY 301 CTGAAGAACATCGTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
DB 301 CTGAAGAACATCGTGAGGGGCAAAACGCTGAGAGTCCGACCTATGATTTTGTGACAC 360
QY 361 TCAAGTTTACAGAGACCAAGGTGTCTACCTTCGAGCGGTGTCTGTTGAGGGATC 420
DB 361 TCAAGTTTACAGAGACCAAGGTGTCTACCTTCGAGCGGTGTCTGTTGAGGGATC 420
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
DB 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCCTCTTCTGTGACACC 480
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGTTCTCCGGGACGTGCGCCGAGGGAGGACCTG 540
DB 481 GACTCCGACGTCAAGCTGTCTCGAAGTTCTCCGGGACGTGCGCCGAGGGAGGACCTG 540
QY 541 GAGGAGATTTCAGAGGAGTACACACCTTCGTGAAGCGCGCCTTCGAGAGTTCTGCTG 600
DB 541 GAGGAGATTTCAGAGGAGTACACACCTTCGTGAAGCGCGCCTTCGAGAGTTCTGCTG 600
QY 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
DB 601 CCGACAAAGAGTATGCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
QY 661 AACCTGATCGTGACACATCCAGACATTTGTAATGTTGACATCTGCAATGCGACCGA 720
DB 661 AACCTGATCGTGACACATCCAGACATTTGTAATGTTGACATCTGCAATGCGACCGA 720
QY 721 GGAAGGTTCATGAGCGGAGCTACAGAGCGACCTTTCTGAGCGAGGGGACACCTGCGG 780
DB 721 GGAAGGTTCATGAGCGGAGCTACAGAGCGACCTTTCTGAGCGAGGGGACACCTGCGG 780
QY 781 ATGCTGACCTCTGGCAAAAGGTTCATTTGAGAGTTCAGACAGACCCCACTGA 834
DB 781 ATGCTGACCTCTGGCAAAAGGTTCATTTGAGAGTTCAGACAGACCCCACTGA 834

RESULT 2
AX449217
LOCUS AX449217 1624 bp DNA linear PAT 03-JUL-2002

DEFINITION Sequence 1 from Patent WO0202761.
ACCESSION AX449217
VERSION AX449217.1 GI:21697994
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Glucksmann, M.A.
TITLE 57658, a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 1 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
CDS
1..1624
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
94..927
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37771.1"
/db_xref="GI:21697995"
/translation="MASAGEDCESPAPADRPHORPLIGVSGTASGSTVCEKIM
ELIGONEBOROKXVILISODRFYKVLTEQAKALKGQYFNPDHPAPNDLMDRLTK
NIVEGTVFVPTDPTVHSRLPETTVVPAVDVLFEGLIVFSOEIRDMFHLRPVDT
DSVRLSRVLRDVRGRDLEOILTYOTTFVAPAPPEFCLPKKYADVILIPGVNMY
ALNLIYOHIDILINDICMHRGSGNSRYKTFSEBPDHPMLISGKASHLESSRP
H"

BASE COUNT 354 a 427 c 506 g 337 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 1624;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGGGAGGAGAACTCGAGAGCCCCCGCGGAGGCGCAACCTCCGAC 60
DB 94 ATGGCTTCGGGGGAGGAGAACTCGAGAGCCCCCGCGGAGGCGCAACCTCCGAC 153
QY 61 CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGT 120
DB 154 CAGCGGCCCTTCTCGATAGGGGTGAGCGCGGCACTGCGACGGGAAAGTCGACCGTGTGT 213
QY 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
DB 121 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 180
QY 214 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 273
DB 214 GAGAGATCATGAGTTGCTGGGACAGAACGAGGTGAAACGCGGACGCGAAGGTGTCT 273
QY 181 ATCTGAGCCGAGACAGGTTCTCAAGAGTCTGACGCGAGACGAGAGCCAAAGCCTTG 240
DB 274 ATCTGAGCCGAGACAGGTTCTCAAGAGTCTGACGCGAGACGAGAGCCAAAGCCTTG 333
QY 241 AAAGGACATCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGAGCT 300
DB 334 AAAGGACATCAATTTTGAACATCCAGATCCCTTGTATATGATTTGATGACAGAGCT 393
QY 301 CTGAAGAACATCGTGAGGGGCAAAACGCTGAGAGTCCGACCTATATGTTTGTGACAC 360
DB 394 CTGAAGAACATCGTGAGGGGCAAAACGCTGAGAGTCCGACCTATATGTTTGTGACAC 453
QY 361 TCAAGTTTACAGAGACCAAGGTGTCTACCTTCGAGAGCGTGTCTGTTGAGGGCATC 420
DB 454 TCAAGTTTACAGAGACCAAGGTGTCTACCTTCGAGAGCGTGTCTGTTGAGGGCATC 513
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCACCTGCGCTCTTCTGTGACACC 480
DB 514 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCACCTGCGCTCTTCTGTGACACC 573
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCCGGGAGTGTGCGGAGGAGGACCTG 540
DB 574 GACTCCGACGTCAAGCTGTCTCGAAGATTTCTCCGGGAGTGTGCGGAGGAGGACCTG 633
QY 541 GAGGAGATTTCAGAGGAGTACACACCTTCGTGAAGCGCGCCTTCGAGAGTTCTGCTG 600

Db 634 GAGCAATTCGTGACGAGTACACACCTTGTAAGCCGGCTTCGAGAGTCTGCTG 693
Qy 601 CCGACAAAGAGATGCCGATGATCATGCCACGAGAGTGGACAAATATGTTGCCATC 660
Db 694 CCGACAAAGAGATGCCGATGATCATGCCACGAGAGTGGACAAATATGTTGCCATC 753
Qy 661 AACCTGATCGTGACGACATCCAGACATCTGTAATGTGACATCTGCAATGAGCACCA 720
Db 754 AACCTGATCGTGACGACATCCAGACATCTGTAATGTGACATCTGCAATGAGCACCA 813
Qy 721 GAGGGTCCAAATGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 780
Db 814 GAGGGTCCAAATGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 873
Qy 781 ATGCTGACCTCTGGCAAAACGTCACATTTGGAGTCCAGACGACCCCACTGA 834
Db 874 ATGCTGACCTCTGGCAAAACGTCACATTTGGAGTCCAGACGACCCCACTGA 927

RESULT 3
BD157613 2160 bp DNA linear PAT 17-JAN-2003
LOCUS BD157613
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157613
VERSION BD157613.1 GI:27863371
KEYWORDS JP 2002191363-A/12456.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2160)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HEBLIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/12456
PD 09-JUL-2002 JP 2000280990
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAOBU
PI SAITO,K.
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI MAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES
source
FT CDS Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN

Query Match 100.0%; Score 834; DB 6; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 215 GAGAGATCATGAGATTGCTGGGACAGAACGAGTGGAAACAGCGGACGGGAAGGTGTC 274
Qy 181 ATCTTGAAGCAGGACAGAGTTCCTAACAGTCTCGACCGGACAGACAAAGGCCAAGCCTTG 240
Db 275 ATCTTGAAGCAGGACAGAGTTCCTAACAGTCTCGACCGGACAGACAAAGGCCAAGCCTTG 334
Qy 241 AAAGCAGTACAAATTTGACCATCCAGATGCTTGTATATGATTTGATGACAGAGCT 300
Db 335 AAAGCAGTACAAATTTGACCATCCAGATGCTTGTATATGATTTGATGACAGAGCT 394
Qy 301 CTGAAGACATCGTGGAGGGCAAAAACGATGAGTGCAGACCTATGATTTTGTGACACAC 360
Db 395 CTGAAGACATCGTGGAGGGCAAAAACGATGAGTGCAGACCTATGATTTTGTGACACAC 454
Qy 361 TCAAGATTACAGACACACGTTGCTTACCTCGGAGAGTGTGTTGTTGAGGACATC 420
Db 455 TCAAGATTACAGACACACGTTGCTTACCTCGGAGAGTGTGTTGTTGAGGACATC 514
Qy 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACTGGGCTCTTCGTGACACC 480
Db 515 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACTGGGCTCTTCGTGACACC 574
Qy 481 GACTCCGAGTGAAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGAGGAGCTG 540
Db 575 GACTCCGAGTGAAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCGAGGAGGAGCTG 634
Qy 541 GAGCAGATTCTGACCGACATACACACCTTCGTGAGACCGGCTTCGAGAGATTCTGCTTG 600
Db 635 GAGCAGATTCTGACCGACATACACACCTTCGTGAGACCGGCTTCGAGAGATTCTGCTTG 694
Qy 601 CCGACAAAGAGATGTCGATGATCATCCACGAGAGTGGACAAATATGTTGTCATC 660
Db 695 CCGACAAAGAGATGTCGATGATCATCCACGAGAGTGGACAAATATGTTGTCATC 754
Qy 661 AACCTGATCGTGACGACATCCAGACATCTGTAATGTGACATCTGCAATGGACACCA 720
Db 755 AACCTGATCGTGACGACATCCAGACATCTGTAATGTGACATCTGCAATGGACACCA 814
Qy 721 GAGGGTCCAAATGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 780
Db 815 GAGGGTCCAAATGCGGAGCTTCAAGCGGACCTTTTCTGAGCCAGGAGCAACCTGGG 874
Qy 781 ATGCTGACCTCTGGCAAAACGTCACATTTGGAGTCCAGACGACCCCACTGA 834
Db 875 ATGCTGACCTCTGGCAAAACGTCACATTTGGAGTCCAGACGACCCCACTGA 928

RESULT 4
LOCUS AK022317 2160 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ12255 f1s, clone MAMMA1001476, highly similar to URIDINE KINASE (EC 2.7.1.48).
ACCESSION AK022317
VERSION AK022317.1 GI:10433687
KEYWORDS oligo capping, f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagaitsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Makamatsu,A., Nakamura,Y., Nagahara,K., Masuo,Y., Ninomiya,K. and Iwayanagi,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2160)
Isogai,T. and Otsuki,T.
TITLE Direct Submission

JOURNAL

Submitted (23-AUG-2000) Takao Isegai, Helix Research Institute, Genomics Laboratory/ 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
source

Location/Qualifiers
1..2160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MAMMA101476"
/tissue_type="Mammary gland"
/clone_id="MAMMA1"
/note="Cloning vector: pME18SFL3"
95..928
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14010.1"
/db_xref="GI:10433688"
/translation="MASAGDEDCSPAPADRPHPRLIGVSGTASGKSTCEKIM
ELLGQNEVEORORRVVILSDRFYKVLTAQKAKALKGQYNFHPDPAFNDLHRTLK
NIVEGKTEVPYDFTVHSRLPETTVVPAADVLFESILVFSQELRDMHLRLFVDT
DSVYRLSRVLRDVRGRDLBOILQYTTFFVPAPEERCLPTKKYADVILPPGVDMV
AINTLVQHLQDILNDGICKMHRGSGNSRYKTFSEPDHPMLTSGKSHLESSSRP
H"

BASE COUNT 457 a 591 c 671 g 441 t
ORIGIN

Query Match 100.0%; Score 834; DB 9; Length 2160;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGCGACCGTCCGAC 60
95 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGCGACCGTCCGAC 154
61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCGCACTCGACGGGAAGTGCACCGTGTGT 120
155 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCGCACTCGACGGGAAGTGCACCGTGTGT 214
121 GAGAAATCATGAGTCTCTGGGACAGAACGAGGTGAAACGCGCAGCGGAAGTGTCT 180
215 GAGAAATCATGAGTCTCTGGGACAGAACGAGGTGAAACGCGCAGCGGAAGTGTCT 274
181 ATCTGAGCCAGGACAGGTTCTTCAAGGTCTCTGACGCGACAGGACGAAAGCCCTTG 240
275 ATCTGAGCCAGGACAGGTTCTTCAAGGTCTCTGACGCGACGAAAGCCCTTG 334
241 AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTTGAATATGATTTGATGACAGGACT 300
335 AAAGGACAGTCAATTTTGAACCATCCAGATGCTTTTGAATATGATTTGATGACAGGACT 394
301 CTGAAGAACATGTTGAGAGGCAAAACGTTGAGGTGCGACCTATGATTTTGTGACACAC 360
395 CTGAAGAACATGTTGAGAGGCAAAACGTTGAGGTGCGACCTATGATTTTGTGACACAC 454
361 TCAGAGTTTACAGAGACCAAGGTGTCTACCTCTCGGACGAGGTCTTCTTTGAGGAGATC 420
455 TCAGAGTTTACAGAGACCAAGGTGTCTACCTCTCGGACGAGGTCTTCTTTGAGGAGATC 514
421 TTGGTGTCTTCAACGCGAGAGATCCGGGACATGTTTCAACCTGCGCCTCTTCTGTGACACC 480
515 TTGGTGTCTTCAACGCGAGAGATCCGGGACATGTTTCAACCTGCGCCTCTTCTGTGACACC 574
481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGTCGCGACGAGGAGGAGCCTG 540
575 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGTCGCGACGAGGAGGAGCCTG 634

QY 541 GAGCAGATTCTGACGCGAGTACACCACTCTTGTAAGCCGGCCTTGAGGAGTGTCTGCTG 600
DB 635 GAGCAGATTCTGACGCGAGTACACCACTCTTGTAAGCCGGCCTTGAGGAGTGTCTGCTG 694
QY 601 CCGACAAAGAGTATCCGATGATCATCCGAGAGTGGACAAATATGTTGCAATC 660
DB 695 CCGACAAAGAGTATCCGATGATCATCCGAGAGTGGACAAATATGTTGCAATC 754
QY 661 AACCTGATGTCGACACATCCGAGACATTTGTAATGTGTGACATCTTCCAAATGTGACCCA 720
DB 755 AACCTGATGTCGACACATCCGAGACATTTGTAATGTGTGACATCTTCCAAATGTGACCCA 814
QY 721 GAGAGGTCGAATGAGGAGGAGCTTCAAGCGGACCTTTCTGAGCCAGGAGACCACTTGGG 780
DB 815 GAGAGGTCGAATGAGGAGGAGCTTCAAGCGGACCTTTCTGAGCCAGGAGACCACTTGGG 874
QY 781 ATGCTGACCTTGGCAAGCGTGCACATTTGAGTCCAGACGACGACCCCACTGA 834
DB 875 ATGCTGACCTTGGCAAGCGTGCACATTTGAGTCCAGACGACGACCCCACTGA 928

RESULT 5
AF254133 834 bp mRNA linear PRI 02-MAY-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Ho, Y.S. and Johnson, R.K.
Direct Submision
Submitted (10-APR-2000) Oncology Research, Smithkline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA

FEATURES
source

1..834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AK49122.1"
/db_xref="GI:13924750"
/translation="MASAGDEDCSPAPADRPHPRLIGVSGTASGKSTCEKIM
ELLGQNEVEORORRVVILSDRFYKVLTAQKAKALKGQYNFHPDPAFNDLHRTLK
NIVEGKTEVPYDFTVHSRLPETTVVPAADVLFESILVFSQELRDMHLRLFVDT
DSVYRLSRVLRDVRGRDLBOILQYTTFFVPAPEERCLPTKKYADVILPPGVDMV
AINTLVQHLQDILNDGICKMHRGSGNSRYKTFSEPDHPMLTSGKSHLESSSRP
H"

BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN

Query Match 93.9%; Score 783; DB 9; Length 834;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGCGACCGTCCGAC 60
DB 1 ATGGCTTCGGCGGAGCGAAGACTGCGAGAGCCCGCGCGCGACCGTCCGAC 60
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCGCACTCGACGGGAAGTGCACCGTGTGT 120
DB 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCGCACTCGACGGGAAGTGCACCGTGTGT 120


```

QY 601 CCGACAAAGAGTAGTCCGATGTGATCATCCACAGAGAGTGAACAATATGTTGCCATC 660
Db 605 CCGACAAAGAGTAGTCCGATGTGATCATCCGAGAGAGTGAACAATATGTTGCCATC 664
QY 661 AACCTGATCGTGCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGSCACCGA 720
Db 665 AACCTGATCGTGCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGSCACCGA 724
QY 721 GGAGGCTCAATGGGCGGAGCTTCTGAGCCAGGGGACACCCCTGGG 780
Db 725 GGAGGCTCAATGGGCGGAGCTTCTGAGCCAGGGGACACCCCTGGG 784
QY 781 ATGCTGACCTTGGCAACCGTCACTTTGGAGTCCAGACAGACCCCACTGA 834
Db 785 ATGCTGACCTTGGCAACCGTCACTTTGGAGTCCAGACAGACCCCACTGA 838

RESULT 7
BD146824 753 bp DNA linear PAT 17-JAN-2003
LOCUS BD146824
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD146824
VERSION BD146824.1 GI:27852582
KEYWORDS JP 2002191363-A/1667.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 753)
Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saio,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 1667 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/1667
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12O1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
FT source 1..753
FT location/Qualifiers
1..753
/organism="Homo sapiens (human)"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 158 a 196 c 261 g 135 t 3 others

ORIGIN
Query Match 68.9%; Score 575; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 5.3e-292;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCGGGCGGAGCGAAGACTGCGAGAGCCCGCGCGAGCGCCGCGCCGAC 60
Db 95 ATGGCTTCGGGCGGAGCGAAGACTGCGAGAGCCCGCGCGAGCGCCGCGCCGAC 154
QY 61 CAGCGGCCCTTCTCTGATAGGGGTGAGCGGCGCATCTCCAGCGGAAAGTGCACCTGTGT 120
Db 155 CAGCGGCCCTTCTCTGATAGGGGTGAGCGGCGCATCTCCAGCGGAAAGTGCACCTGTGT 214
QY 121 GAGAAATCATGAGTGTCTGGACAGAAAGAGGTGAAACGCGCGACGCGAAAGGTGTC 180
Db 215 GAGAAATCATGAGTGTCTGGACAGAAAGAGGTGAAACGCGCGACGCGAAAGGTGTC 274

```

```

QY 181 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACGGCAGAGGAGGCCAAGCCCTTG 240
Db 275 ATCTGAGCCAGGACAGGTTCTCAAGAGTCTCGACGGCAGAGGAGGCCAAGCCCTTG 334
QY 241 AAAGACAGTACATTTTGCATCCATCCAGATCCCTTGTATGATTTGATGACAGAGCT 300
Db 335 AAAGACAGTACATTTTGCATCCATCCAGATCCCTTGTATGATTTGATGACAGAGCT 394
QY 301 CTGAAGAACTCGTGAAGGGCAAAACGGTGGAGAGTCCGACCTATGATTTTGGACACAC 360
Db 395 CTGAAGAACTCGTGAAGGGCAAAACGGTGGAGAGTCCGACCTATGATTTTGGACACAC 454
QY 361 TCAAGTTTACAGAGACCAAGTGTCTACCTCGCGGACGTTGCTGTTGAGGGCATC 420
Db 455 TCAAGTTTACAGAGACCAAGTGTCTACCTCGCGGACGTTGCTGTTGAGGGCATC 514
QY 421 TTGGTGTCTACAGCCAGAGATCCGGACATGTCACCTGGCCCTCTTCGTGACACC 480
Db 515 TTGGTGTCTACAGCCAGAGATCCGGACATGTCACCTGGCCCTCTTCGTGACACC 574
QY 481 GACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 540
Db 575 GACTCCGACGTCAAGCTGTCTCGAAGAGTTCTCCGGACGTGGCCGAGGAGGACCTG 634
QY 541 GAGGAGTTCTGACGCGAGTACACACCTTGTTGA 575
Db 635 GAGGAGTTCTGACGCGAGTACACACCTTGTTGA 669

RESULT 8
AF125106 1395 bp mRNA linear PRI 07-FEB-2002
LOCUS AF125106
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF125106
VERSION AF125106.1 GI:18568108
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1395)
Kin,Y.R., Yu,L. and Zhao,S.Y.
Cloning of a new human cDNA similar to Mus musculus uridine kinase
mRNA
Unpublished
2 (bases 1 to 1395)
Ding,J.B., Yu,L. and Zhao,S.Y.
Direct Submission
Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of
Genetics, Pudan University, No. 220 Handan Rd., Shanghai 200433,
People's Republic of China
Location/Qualifiers
source 1..1395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
8..838
/codon_start=1
/product="uridine kinase"
/protein_id="AAU75943.1"
/db_xref="GI:18568109"
/translation="MASAGGAGARAGANRPHORPRLIGVGSGTASGKSTVCEKIME
LIGQNEVERBGRKQVILSDRFYKILTAQKALKQYNFPHPAFENDLNRITLKN
IVEGKVEVPTDFTVHSRLPTVTVYADVVLFGLIVFYQELIDMHLRLFVTD
SDVRLSRVLRDVRGRDLEQLITVTFVKPAFEFCLPTKKYADVILPRGVNVA
INLIYVIOHIDILNDICIKWNRGSGNGRTYKTFSEBGDHPMLTSGSKRSLSSSRPH
"

BASE COUNT 325 a 352 c 411 g 307 t

ORIGIN
Query Match 62.0%; Score 517; DB 9; Length 1395;
Best Local Similarity 99.8%; Pred. No. 2.3e-261;

```

Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	172	AGGGGTGATCTCTGAGCCAGGACAGTTTACAAAGTCTCTGACCGAGAGCAAGAGCC	231	
DB	176	AAAGGTGATCTCTGAGCCAGGACAGTTTACAAAGTCTCTGACCGAGAGCAAGAGCC	235	
QY	232	AAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATCCCTTGTGAATATGATTTGATG	291	
DB	236	AAAGGCTTGAAGAGACAGTACATTTTGAACATCCAGATCCCTTGTGAATATGATTTGATG	295	
QY	292	CACAGACCTCTGAAGAACATCTGTGAGGGGCAAAACGGTGGAGGTGCCCATATGATTTT	351	
DB	296	CACAGACCTCTGAAGAACATCTGTGAGGGGCAAAACGGTGGAGGTGCCCATATGATTTT	355	
QY	352	GTGACACACTCAAGGTACAGAGACCGAGGTGTCTACCTTCGAGCGGTGTCTGTTT	411	
DB	356	GTGACACACTCAAGGTACAGAGACCGAGGTGTCTACCTTCGAGCGGTGTCTGTTT	415	
QY	412	GAGGGATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACCTGCCCTCTTC	471	
DB	416	GAGGGATCTTGGTGTCTACAGCCAGAGATCCGGGACATGTTCACCTGCCCTCTTC	475	
QY	472	GTGACACACGACTCCGACGTCTGCTGTCTGAAGAGTTCTCCGGGACGTGCGCCGAGG	531	
DB	476	GTGACACACGACTCCGACGTCTGCTGTCTGAAGAGTTCTCCGGGACGTGCGCCGAGG	535	
QY	532	AGGGAAGCTGGAAGAGATTCTGACCGAGTACCACTTCTGTGAAGCCGGCTTCAGAGAG	591	
DB	536	AGGGAAGCTGGAAGAGATTCTGACCGAGTACCACTTCTGTGAAGCCGGCTTCAGAGAG	595	
QY	592	TTCTGCTGCGGCAAAAGATATGCGATGTATCATCCACAGAGAGTGAATATG	651	
DB	596	TTCTGCTGCGGCAAAAGATATGCGATGTATCATCCACAGAGAGTGAATATG	655	
QY	652	GTTCGATCAATCACTGATCTGTGACGACATTCAGAGCAATTCGATGATGATCTGCAAA	711	
DB	656	GTTCGATCAATCACTGATCTGTGACGACATTCAGAGCAATTCGATGATGATCTGCAAA	715	
QY	712	TGGCACCGAGAGAGTCCATGGGCGGA	739	
DB	716	TGGCACCGAGAGAGTCCATGGGCGGA	743	
RESULT 9				
BC015547				
LOCUS				
DEFINITION				
Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9668				
ACCESSION				
BC015547				
VERSION				
BC015547.1 GI:15930229				
KEYWORDS				
MGC.				
SOURCE				
Homo sapiens (human)				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE				
1 (bases 1 to 2072)				
Strausberg, R.				
Direct Submission				
Submitted (01-OCT-2001) National Institutes of Health, Mammalian				
Gene Collection (MGC), Cancer Genomics Office, National Cancer				
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
USA				
NHL-MGC Project URL: http://mgc.nci.nih.gov				
Contact: MGC help desk				
Email: cgabbs-remail.nih.gov				
Tissue Procurement: ATCC				
CDNA Library Preparation: Life Technologies, Inc.				
DNA Sequencing by: Sequencing Group at the Stanford Human Genome				
Center, Stanford University School of Medicine, Stanford, CA 94305				
Web site: http://www-shgc.stanford.edu				
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,				
REMARK				
COMMENT				

R. M.				
Clone distribution: MGC clone distribution information can be found				
through the I.M.A.G.E. Consortium/BLND at: http://image.llnl.gov				
Series: IRK Plate: 20 Row: 5 Column: 5				
This clone was selected for full length sequencing because it				
passed the following selection criteria: matched mRNA gi: 1389252.				
FEATURES				
source				
1..2072				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="MGC:9668 IMAGE:3845821"				
/tissue_type="Colon, adenocarcinoma"				
/clone_id="NIH_MGC_65"				
/lab_host="DH10B"				
/note="Vector: PCMV-SPORT6"				
CDS				
87..692				
/codon_start=1				
/product="similar to uridine-cytidine kinase 1"				
/protein_id="AAH15547.1"				
/db_xref="GI:15930230"				
/translation="MASAGEDCESPAPRPHQRPLLIGVSGTASGKTVCEKIM				
ELLGONEVEORORRVVILSDRFYVLTAEQAKALKQYNFDPHPDAFNDLMLRTLK				
NIVEGTVFPTDVFVTRRLPBTIVVPADVVLVEGLIVFSQEIADWFHLRLFVDT				
DSDVLSRRDKVCRCDHPTRSQVGHQPDAAHPHSEW"				
BASE COUNT				
455 a 559 c 637 g 421 t				
ORIGIN				
Query Match				
Best Local Similarity 100.0%; Pred. No. 1.3e+25; Length 2072;				
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGGCTTGGCGGAGAGCGAAGACTGCGAGACCCCGCGAGCGGACCGTCCGAC	60	
DB	87	ATGGCTTGGCGGAGAGCGAAGACTGCGAGACCCCGCGAGCGGACCGTCCGAC	146	
QY	61	CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAGATCGACCGGTGT	120	
DB	147	CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAGATCGACCGGTGT	206	
QY	121	GAGAAGATCATGAGGTGCTGGGACAGAACGAGGTGAAACAGCGGACGGAAGGTGTC	180	
DB	207	GAGAAGATCATGAGGTGCTGGGACAGAACGAGGTGAAACAGCGGACGGAAGGTGTC	266	
QY	181	ATCTGAGCCAGAGACAGTCTTCAAGAGTCTGACGCGAGACAGAAAGGCCAAGGCTTG	240	
DB	267	ATCTGAGCCAGAGACAGTCTTCAAGAGTCTGACGCGAGACAGAAAGGCCAAGGCTTG	326	
QY	241	AAAGACATTAATAATTTTCAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACT	300	
DB	327	AAAGACATTAATAATTTTCAACATCCAGATGCTTGTGATTAATGATTTGATGACAGACT	386	
QY	301	CTGAAGAACATGTTGAGAGGCAAAACGAGTGAAGTGCAGCTTATGATTTTGTGACACAC	360	
DB	387	CTGAAGAACATGTTGAGAGGCAAAACGAGTGAAGTGCAGCTTATGATTTTGTGACACAC	446	
QY	361	TGAAGTTTCAAGAGACACAGGTGTCTACCTTGCAGAGGTGTCTGTTTGAAGGCGATC	420	
DB	447	TGAAGTTTCAAGAGACACAGGTGTCTACCTTGCAGAGGTGTCTGTTTGAAGGCGATC	506	
QY	421	TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACACAC	480	
DB	507	TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTGCGCTCTTGTGACACAC	566	
QY	481	GACTCCGACGTGAGGCTGTCTGAAAGAG	508	
DB	567	GACTCCGACGTGAGGCTGTCTGAAAGAG	594	
RESULT 10				
AK057848				
LOCUS				
AK057848				
2228 bp				
mRNA				
linear				
PRI 31-OCT-2001				

DEFINITION Homo sapiens cDNA FLJ25119 fis, clone CBR05878, highly similar to
UNIDISE KINMSE (EC 2.7.1.48).
ACCESSION AK057848
VERSION AK057848.1 GI:16531809
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiya,M., Suzuki,Y., Hata,H.,
Nakagawa,K., Mizuno,S., Morinaga,M., Kawamura,T., Sugiyama,T.,
Irie,R., Otsuki,T., Sato,H., Nishikawa,T., Sugiyama,A.,
Kwakami,B., Nagai,K., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2228)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.
FEATURES
source location/Qualifiers
1..2228
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR05878"
/tissue_type="brain"
/clone_lib="CBR"
/note="Cloning vector: pME18SFL3"
BASE COUNT 478 a 598 c 696 g 456 t
ORIGIN
Query Match 59.6%; Score 497; DB 9; Length 2228;
Best Local Similarity 100.0%; Pred. No. 8.1e-251;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 107 AGTCGACCGTGTGTGAGAGATCATGTAGTGTGGACAGAAAGAGTGAACAGCGGC 166
DB 212 AGTGAACGCTGTGTGAGAGATCATGTAGTGTGGACAGAAAGAGTGAACAGCGGC 271
QY 167 AGCGGAAGGTGTATCTTGAAGCAGAGACAGGTTCTCAAGAGTCTTGAACGCGCAGAGAGA 226
DB 272 AGCGGAAGGTGTATCTTGAAGCAGAGACAGGTTCTCAAGAGTCTTGAACGCGCAGAGAGA 331
QY 227 AGGCGAAGGCTTGAAGGACAGTACAAATTTGACATCCAGATGCTTGAATATATAT 286
DB 332 AGGCGAAGGCTTGAAGGACAGTACAAATTTGACATCCAGATGCTTGAATATATATAT 391
QY 287 TGATGACAGAGACTCTGAGAAACATCGTGAAGGCAAAAGCGTGAAGAGTCCGACCTATG 346
DB 392 TGATGACAGAGACTCTGAGAAACATCGTGAAGGCAAAAGCGTGAAGAGTCCGACCTATG 451
QY 347 ATTTTGTGACACTCAAGGTTACAGAGACACGCTGTCTACCTTGCAGAGCTGTTC 406
DB 452 ATTTTGTGACACTCAAGGTTACAGAGACACGCTGTCTACCTTGCAGAGCTGTTC 511
QY 407 TGTGTGAGGAGCTTGTGTGTCTACAGCCAGAGAGATCCGGAGATGTTCACTCTGCGCC 466
DB 512 TGTGTGAGGAGCTTGTGTGTCTACAGCCAGAGAGATCCGGAGATGTTCACTCTGCGCC 571
QY 467 TCTTGTGACACCGACTCCGACGTGAGCTGTCTCAAGAGTCTCCGGAGAGCTGCGCC 526

DB 572 TCTTGTGACACCGACTCCGACGTGAGCTGTCTGAGAGATTCCTCCGGAGAGTCCGCC 631
QY 527 GAGGAGAGGACCTGAGAGAGATTTCTGACGATACACCACTTCTGTAAGCCGCTTCG 586
DB 632 GAGGAGAGGACCTGAGAGAGATTTCTGACGATACACCACTTCTGTAAGCCGCTTCG 691
QY 587 AGGAGTTCTGCGCTGCGC 603
DB 692 AGGAGTTCTGCGCTGCGC 708
RESULT 11
AX540411 734 bp DNA linear PAT 23-NOV-2002
LOCUS AX540411
DEFINITION Sequence 23 from Patent WO02055738.
ACCESSION AX540411
VERSION AX540411.1 GI:25273437
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE
AUTHORS Panzer,S.R., Lincoln,S.E., Altus,C.M., Dufour,G.E., Hillman,D.L.,
Jones,A.L., Lam,T.C., Liu,T.F., Harris,B., Flores,V., Daffo,A.,
Marwaha,R., Chen,A.U., Chang,S.C., Gestlin,E.H., Peralta,C.H.,
David,M.H. and Lewis,S.A.
TITLE Molecules for disease detection and treatment
JOURNAL Patent: WO 02055738-A 23 18-Jul-2002;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: LI:235557.12:2001JAN12"
BASE COUNT 166 a 181 c 226 g 154 t 7 others
ORIGIN
Query Match 40.5%; Score 338; DB 6; Length 734;
Best Local Similarity 100.0%; Pred. No. 8.6e-167;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 266 CAGATGCTTGTGATATGATTTGATGACAGACTCTGAAGAACATCGTGAAGGGA 325
DB 12 CAGATGCTTGTGATATGATTTGATGACAGACTCTGAAGAACATCGTGAAGGGA 71
QY 326 CGGTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGTGG 385
DB 72 CGGTGAGAGTGCAGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGTGG 431
QY 386 TCTTACCTGCGGACGTGTGTCTGTTTGAAGGACATCTTGTGTCTTACAGCCAGAGATCC 445
DB 132 TCTTACCTGCGGACGTGTGTCTGTTTGAAGGACATCTTGTGTCTTACAGCCAGAGATCC 491
QY 446 GGGACATGTTTCCACCTGCGGCTTCTGTTGAGACACGACCTCCAGCTCAGGCTGTCCGA 505
DB 192 GGGACATGTTTCCACCTGCGGCTTCTGTTGAGACACGACCTCCAGCTCAGGCTGTCCGA 551
QY 506 GAGTTCTCCGAGAGCTGCGCCGAGAGAGGAGACTGAGAGAGATTCTGAGCAGTACCA 565
DB 252 GAGTTCTCCGAGAGCTGCGCCGAGAGAGGAGACTGAGAGAGATTCTGAGCAGTACCA 611
QY 566 CTTTGTGAAAGCCGCTTCTGAGAGAGTCTGCTGCGC 603
DB 312 CTTTGTGAAAGCCGCTTCTGAGAGAGTCTGCTGCGC 349
RESULT 12
AL358781 147492 bp DNA linear PRI 06-OCT-2001
LOCUS AL358781/c
DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete

sequence.
 AL358781
 VERSION AL358781.19 GI:13751418
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 147492)
 Corby,N.
 Direct Submission
 Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Apr 21, 2001 this sequence version replaced gi:13395549.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-33406 is from the library RPCT-11.2 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6
 This sequence is the entire insert of clone RP11-33406 The true left end of clone RP11-4007 is at 113870 in this sequence. The true right end of clone RP11-643B14 is at 63282 in this sequence.
 Location/Qualifiers
 1..147492
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-33406"
 /clone_11b="RPCT-11.2"
 30143..30199
 /note="Sequence from overlapping clone RP11-643B14 (AL354855). Assembly confirmed by restriction digest."
 misc_feature
 BASE COUNT 32592 a 36290 c 37577 g 40633 t
 ORIGIN
 Query Match 21.9%; Score 183; DB 9; Length 147492;
 Best Local Similarity 100.0%; Pred.No. 4.8e-85;
 Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 652 GTTGCCATCACTGATCGGAGACATCCAGACATTTCTGAATGGTGACATCTGCAA 711
 Db 129488 GTTGCCATCACTGATCGGAGACATCCAGACATTTCTGAATGGTGACATCTGCAA 129429
 QY 712 TGGCAGCGAGAGGCTCAATGGCGAGATCAAGGAGACTTTCTGAGCCAGGGAG 771
 Db 129428 TGGCAGCGAGAGGCTCAATGGCGAGATCAAGGAGACTTTCTGAGCCAGGGAG 129369
 QY 772 CACCTGGAGTGTGACCTGTGGCAACGGTCACTATTGGATGCCAGACAGACCCAC 831
 Db 129368 CACCTGGAGTGTGACCTGTGGCAACGGTCACTATTGGATGCCAGACAGACCCAC 129309

QY 832 TGA 834
 Db 129308 TGA 129306
 RESULT 13
 BV077668/c
 LOCUS
 DEFINITION S212P6227FD8.T0 CZECHII/E1 Mus musculus STS genomic, sequence
 ACCESSION BV077668
 VERSION BV077668.1 GI:31193463
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 660)
 Made,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C., Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
 The mosaic structure of variation in the laboratory mouse genome
 Nature 420 (6915), 574-578 (2002)
 22354684
 12468852
 COMMENT
 CONTACT: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kerstin@genome.wi.mit.edu
 Primer A: None
 Primer B: None
 STS size: 660
 STS protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SYRMJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGS/Cv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.
 Location/Qualifiers
 1..660
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="CZECHII/E1"
 /db_xref="taxon:10090"
 /map="x 2 22-615 32611326-32611918"
 /clone_11b="CZECHII/E1"
 <1..660
 STS
 BASE COUNT 174 a 196 c 174 g 116 t
 ORIGIN
 Query Match 4.3%; Score 36; DB 11; Length 660;
 Best Local Similarity 100.0%; Pred.No. 3e-07;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 435 CCAGAGATCCGGAGCATGTCCACCTGGCCCTCTT 470
 Db 435 CCAGAGATCCGGAGCATGTCCACCTGGCCCTCTT 400
 RESULT 14
 MUSURKI
 LOCUS MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
 DEFINITION Mus musculus uridine kinase mRNA, partial cds.
 ACCESSION L31783
 VERSION L31783.1 GI:471980
 KEYWORDS

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1810)
JOURNAL The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites
MEDLINE Eur. J. Biochem. 222 (1), 9-19 (1994)
PUBMED 94259063
REFERENCE 8200357
AUTHORS 2 (bases 1 to 1810)
TITLE Ropp, P.A. and Traut, T.W.
JOURNAL Cloning and expression of a cDNA encoding uridine kinase from mouse
MEDLINE Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
PUBMED 8951040
FEATURES
CDS
1. 1810
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="Brain"
/EC_number="2.7.1.48"
/standard_name="uridine-cytidine kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AA50568.1"
/db_xref="GI:471981"
/translation="RPPRPPLIGVSGTASGKSTVCEKIMELLGQNEVDPRORLKI
LSQCFKVLTAQKAKLQGVNFDHDAFNDMLKNIIVGKTVFVPTDYDT
HSRLPRTVVPADVLEGLVFTYDIDMFLRLFPDSDPRLSRVLDVVG
RDLEQILQVTAFTKPAPEEPCPTKADVIIRGVDMVAIQLQIIDLNGDL
CKRRGPGNRRHRTPEPDHGVATGRSHLESSSRP"
misc_binding
Join(37, 63, 412, 426, 451, 465)
/bound_moiety="ATP at catalytic site"
polya_signal
1790..1795
BASE COUNT 411 a 482 c 512 g 405 t
ORIGIN
Query Match 4.3%; Score 36; DB 10; Length 1810;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

435 CCAGGAGATCCGGACATGTCACCTGCGCCTTT 470
Db 384 CCAGGAGATCCGGACATGTCACCTGCGCCTTT 419

RESULT 15
BC025146 1959 bp mRNA linear ROD 16-APR-2003
LOCUS Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
DEFINITION MGC:36231 IMAGE:4913412), complete cds.
ACCESSION BC025146
VERSION BC025146.1 GI:19263563
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1959)
JOURNAL Strausberg, R.L., Pelting, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usadi, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
DIRECT Submission
Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Guanarato, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsged, H.,
Kowit, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>
Series: IRAC Plate: 61 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 1959
/organism="Mus musculus"
/mol_type="mRNA"
/ecr="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36231 IMAGE:4913412"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. 1959
/gene="Umpk"
/db_xref="LocusID:22245"
/db_xref="MGI:98904"
57..890
/codon_start=1
/product="Umpk protein"
/protein_id="AAH25146.1"
/db_xref="GI:19263564"
/db_xref="LocusID:22245"
/translation="MASAGGSGSBSAARADPPRPPLIGVSGTASGKSTVCEKIM
ELLGQNEVDPRORLKI
LSQCFKVLTAQKAKLQGVNFDHDAFNDMLKNIIVGKTVFVPTDYDT
HSRLPRTVVPADVLEGLVFTYDIDMFLRLFPDSDPRLSRVLDVVG
RDLEQILQVTAFTKPAPEEPCPTKADVIIRGVDMVAIQLQIIDLNGDL
CKRRGPGNRRHRTPEPDHGVATGRSHLESSSRP"
H"
BASE COUNT 466 a 511 c 561 g 421 t
ORIGIN
Query Match 4.3%; Score 36; DB 10; Length 1959;

Wed Nov 26 18:16:25 2003

us-09-896-522-3.01i.rge

Page 11

Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	435	CCAGAGATCCGGGACATGTTCACCTGCgcctcTT	470
Dd	491	CCAGAGATCCGGGACATGTTCACCTGCgcctcTT	526

Search completed: November 25, 2003, 05:12:22
Job time : 3245.82 secs

This Page Blank (uspto)


```

Pred. No.: 1.26e-37 Length: 277
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.25% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK1_MOUSE (1-277)

QY 298 ACCTGAGAGACATCGGAGGGGAAACGCTGAGAGTCCGACCTTATGATTTTGACCA 357
    |||
Db 100 ThrluylasnlleValgluglylYsrhValglValProtnrTYraSpPheValThr 119

QY 358 CACTCAAGTTACAGACAGACGAGTCTACCCCTCGAGACGTTCTGTTTGAGAGGC 417
    |||
Db 120 HlserrArgleuProglunrThValValTYrProAlaSpValValLeupheglugly 139

QY 418 ATCTTGCTGTTCTAC 432
    |||
Db 140 lleleuValpethyr 144

RESULT 3
UCK2_HUMAN
ID UCK2_HUMAN STANDARD; PRT; 261 AA.
AC Q9BZX2; Q96KG5;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johanson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases.";
RN Mol. Pharmacol. 59:1181-1186(2001).
[2]
RP SEQUENCE OF 15-261 FROM N.A.
RX TISSUE= fibrosarcoma;
RA MEDLINE=21385121; PubMed=11494055;
RA Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RA Fukushima M.;
RT "Cloning and expression of uridine/cytidine kinase cDNA from human
RT fibrosarcoma cells.";
RL Int. J. Mol. Med. 8:273-278(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, and N(4)-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Expressed in placenta.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AF236637; AAK14053.1; -.
DR EMBL; AB062451; BAB56162.1; -.
DR Genbank; H0NC12562; UMPK.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Urid_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00478; PHRI1KINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding.
FT NE BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;

Alignment Scores:
Pred. No.: 4.59e-17 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK2_HUMAN (1-261)

QY 565 ACCTGAGAGCGGCGCTTCGAGAGGTTCTGCTGCCGACAAAGATGCGGATGNG 624
    |||
Db 187 ThpPheVallyProAlaPhegluglyupheCysleuProtnrThylsTYrAlaSpVal 206

QY 625 ATCATCCACGAGGA 639
    |||
Db 207 llelleProArggly 211

RESULT 4
UCK2_MOUSE
ID UCK2_MOUSE STANDARD; PRT; 261 AA.
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Last Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johanson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL human uridine-cytidine kinases.";
RN Mol. Pharmacol. 59:1181-1186(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, and N(4)-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: AF236636; AAK14052.1; -
DR MGI: 1931744; Uck2.
DR InterPro: IPR006082; PRK.
DR InterPro: IPR006083; PRK URK.
DR InterPro: IPR00764; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS: PR00478; PHRIBKINASE.
DR PRINTS: PR00988; URIDINKINASE.
DR TIGRFAMs: TIGR00235; uck; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;

Alignment Scores:
Pred. No.: 4,59e-17 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK2_MOUSE (1-261)

QY 565 ACCTTCGAGAGCGGCTTGGCTGCGGACAGAGATGCGGATGTG 624
Db 187 ThreVallylPProAlaValIleuPheGluGluPheCysLeuProIlylValAlaSpVal 206

QY 625 ATCATCCGACGAGGA 639
Db 207 IleIleProArgIly 211

RESULT 5
ID UCK_DROME STANDARD; PRK; 260 AA.
AC Q9VC99;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC SMPAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter S.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskearn D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palenik K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA SVIRKAS R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003747; AAF56274.1; -
DR FLYBase: FBgn0039179; CG6364.
DR InterPro: IPR006083; PRK URK.
DR InterPro: IPR00764; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS: PR00988; URIDINKINASE.
KM Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2EE57423704925B6 CRC64;

Alignment Scores:
Pred. No.: 8.73e-07 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-3 (1-834) x UCK_DROME (1-260)

QY 388 TACCCGAGAGCGGCTTGGCTTGGAGGACATGTTGTTCTAC 432
Db 135 TyPProAlaSpValIleuPheGluGluIlyleValAlaSpVal 149

Search completed: November 25, 2003, 07:54:03
Job time : 10.6522 secs

```


Db 241 GlycylserAenGlyArgSerTyrIleThrPheSerGluProGlyAspHisProGly 260

RESULT 2

US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:

Pred. No.:	2,21e-35	Length:	277
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.25%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-3 (1-834) x US-09-536-647-3 (1-277)

QY	298	ACTGTGAAGACATCGTGGAGGCAAA	CGGTGAGGTCGACCTATGATTTGTGCA	357
Db	100	ThrIeuLysasnIleValGluGlyLysThrValGluValProtnTyrAspPheValThr		119
QY	358	CACCTAAGGTATACAGAGACCA	CGGTGCTTACCTCGCGACGTGTTGAGGAC	417
Db	120	HisSerArgIeuProGluThrThrValValTyrProIleAspValIleuPheGluGly		139
QY	418	ATCTTGTTGTTCTTAC	432	
Db	140	IleIeuValPheTyr	144	

Search completed: November 25, 2003, 08:01:01
Job time : 9.99146 secs

Db 161 AsperasperValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
Qy 541 GAGCAGATTCTGACGCGATACACCACTTCTGTAAGCCGGCTTCGAGAGTTCTGCTG 600
Db 181 GluInIleLeuThrGlnTyrThrThrPheValLysProIaPheGluGluPheCysLeu 200
Qy 601 CCGCAAGAAGTATGCGCATGTATCATCCACGAGAGAGGACAATATGTTGGCCATC 660
Db 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyAlaAspMetValAlaIle 220
Qy 661 AACCTGATCTGCAGACATCATCAGACATTCGATGATGTGACATCTGCMAATGGCAGCA 720
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrIleArg 240
Qy 721 GGAGGCTTCATGGCGGAGCTACACAGCGACCTTTTGTAGCCAGGGGACACCTGGG 780
Db 241 GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
Qy 781 ATGCTGACCTCTGGCAACGGTCACTTTGGAGTCCAGCAGACCCAC 831
Db 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
RESULT 2
US-10-029-386-28765
Sequence 28765, Application US/10029386
Publication No.: US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: ABOICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 28765
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9, 1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8.00e-26
US-10-029-386-28765
Alignment Scores:
Pred. No.: 6.02e-47 Length: 60
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 12 Gaps: 0
US-09-896-522-3 (1-834) x US-10-029-386-28765 (1-60)
Qy 652 GTTGCATCAACCTGATGTGACAGACATCCAGACATTTGAATGTGACATTCGAAA 711
Db 1 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 20
Qy 712 TGGCAGCAGAGAGGTCCAAATGGGGAGCTACAAAGCGGACCTTTCTGACCGAGGAGC 771
Db 21 TrpHisArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 40
Qy 772 CACCTGGAGATGTGACCTCTGGCAACGGTCACTTTGAGTCCAGCAGACAGACCCAC 831

Db 41 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 60
RESULT 3
US-09-925-300-1160
Sequence 1160, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruden,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1160
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (38)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (46)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (155)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (169)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160
Alignment Scores:
Pred. No.: 2.05e-14 Length: 337
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.03% Indels: 0
DB: 10 Gaps: 0
US-09-896-522-3 (1-834) x US-09-925-300-1160 (1-337)
Qy 565 ACCTGTGTAAGCGCGCTTCGAGAGATTCTGCTGCCGCAAGAAGTATGCGCATGTG 624
Db 263 ThrPheValLysProIaPheGluGluPheCysLeuProThrLysTyrAlaAspVal 282
Qy 625 ATCATCCACGAGAGA 639
Db 283 IleIleProArgGly 287
RESULT 4
US-09-896-522-5
Sequence 5, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5

; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Alignment Scores:

Pred. No.:	4,83e-05	Length:	125
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.42%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-3 (1-834) x US-09-896-522-5 (1-125)

QY 736 CGAGCTACAAGCGGACCTTTCTGAGCCAGGGAGCCACCCCTGGG 780
|||
Db 94 ArgSerTrpLysArgThrPheSerGluProGlyAspHisProGly 108
|||

Search completed: November 25, 2003, 08:04:08
Job time : 30.1798 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:43:14 ; Search time 15.7775 Seconds
(without alignments)
10166.994 Million cell updates/sec

Title: US-09-896-522-3
Perfect score: 277
Sequence: 1 atgagctcgagcgagcgagcgcga.....ccagcagcagaccaccactga 834

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09896522/runat_21112003_184145_3234/app_query.faeta_1.2830
-DB=PIR_76 -QFMT=faetan -SUFFIX=n2p_ol1.rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MAXLEN=200000000
-USER=US09896522.ccn_1_1_68.rnat_21112003_184145_3234 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match length DB ID Description

No matches found

Search completed: November 25, 2003, 07:59:52
Job time : 15.7775 secs

This Page Blank (uspto)

XX	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AAM39842-AAM42213) with neotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localized neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilization of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemoblastic/chemokine activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemia and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
CC	
XX	
SQ	Sequence 277 AA;
 Alignment Scores:	
Pred. No.:	5, 11e-272
Score:	277.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	22
	Gaps: 0
 US-09-896-522-3 (1-834) x AAM39502 (1-277)	
OY	1 ATGGCTTCGGCGGAGGCGGAACACTGCCAGACCCTCGGAGCGCAACCGTCCGAC 60
Db	1 MeelalaberrlaaglyglnumprCyvGuisetProalaProglualaaPrpionhis 20
OY	61 CAGCGGCCTTCCTGTATAGGGGTGACCGCGGCGCATCGCACGCGGGAAGTCCACCGTGNT 120
Db	21 GlnhrgrProheulleeglyvalserGylgtlrhlaseGilylssetThlValCys 40
OY	121 GAAGAATCATGAGATTGCTGGGACAAGAACAGGTGGAAACAGCGGCACGGAAGGTGTC 180
Db	41 GlutyllelmetGlueulueuglGlnanngluValGluGlnatrgGlnatrgValVal 60
OY	181 ATCTGAGCGAGGAKOGAGTTCTTAACAAGTCTCGAGGACGAGACAGAAAGCCAAAGCCTTG 240
Db	61 lleuseneterlmprkPrrhetYtLysValleuthrlaagluGlnlyValaleu 80
OY	241 AAAAGCAGTCAAAATTTTGCATCCAGATGCCCTTGATATGATTGATGCACAGACT 300
Db	81 lyeGlyGlnlryrtnphneahsrhsrProdarllaPhedrsenArbleMehIsargThr 100
OY	301 CTGAACAACATCGTGGAGGGCAAACAGGTGAGGTGCGCATATGATTTTGTGACAC 360
Db	101 LeulyasnnlleValGluGlyLysThrValGluValProthrYUAspRheValThnHis 120
OY	361 TCAAGGTTACCAAGACCAAGGAGGTCTACCTCGGACAGTGGATTCTGTTGAGGGCATC 420
Db	121 SethrigleurProdlurthrValValIlyrrrolaaRerValValLeuhrebeluglyile 140
OY	421 TTGGTGTCTTACACGACAGAGATCCGGGACATGTTCCACTGCGCTTCTGTGGACAC 480
Db	141 leuValaPrrhetYserGingluilleatrgdsrmertPrenhIsleuatrgLeuRheValAspThr 160
OY	481 GAGTCGAGGTCAAGGCTGTCTGAAAGTTCACGGGACGTGCGCGCAAGGGAGGACCTG 540
Db	161 AspssetarPaallaTrguenezeradgaValleuhtdgarPaValdaVzglyUAgsRleu 180
OY	541 GAGCAGATTGTGACGAGTACACACACTTCGTAAGCGGCGCTTCGAGAGATTTCGCTG 600
Db	181 GluGlnlleleuthrGlnlrytThrThrPheValValysprollaPhedgluRheCyleu 200
OY	601 CCAGCAAAAAGATATGCGCATGTGATCATCCCAAGAGGTGGACATAATGTTGGCATC 660
Db	201 ProthrlyslsyrYrllaaSpralllelleProldrglyValaaAspmeeValaalle 220

QY	666	AACTGATGCTGGAGACATCCAGGACATTCGATGGTGCACATCTGCAATAATGCACCGA	720
Db	221	AsnLeuIleValAlaIlnHisIleGlnMetPheIleuAsnGlyAspIleCysGlySerPheAsnG	240
QY	721	GGAGGGGTCCATATGGGCGAGACTTACAAGCGGACCTTTTCTGAGCCAGGAGGACCACTTGGG	780
QY	781	ATGCTGACCTCTGGGCAACGGGTCAATTGGAGTCCACAGACAGACCCAC	831
Db	241	GlyGlySerAsnGlnGlySerGlyTyrIleAsnThrPheSerGlnProGlyAspHisProGly	260
Db	261	MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis	277
RESULT 2			
ID	AAAB93941	standard; Protein; 277 AA.	
XX	AAAB93941;		
AC	26-JUN-2001	(first entry)	
DT	Human protein sequence SEQ ID NO:13952.		
DE	Human, primer; detection; diagnosis; antisense therapy; gene therapy.		
XX	Homo sapiens.		
KW	EPI074617-A2.		
PN	07-FEB-2001.		
PD	28-JUL-2000; 2000EP-0116126.		
PF	29-JUL-1999; 99JP-0248036.		
PR	27-AUG-1999; 99UP-0300253.		
PR	11-JAN-2000; 2000JP-0118776.		
PR	02-MAY-2000; 2000JP-0183767.		
PR	09-JUN-2000; 2000JP-0241899.		
XX	(HELI-) HELIX RES INST.		
PA	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
XX	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
P1	WPI; 2001-318749/34.		
DR	Primer sets for synthesizing polynucleotides, particularly the 5602		
XX	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
XX	full-length cDNAs -		
BS	Claim 8; SEQ ID 13952; 2537bp + CD ROM; English.		
XX	The present invention describes primer sets for synthesizing 5602		
CC	full-length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification, where the		
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
CC	of an oligonucleotide comprising a sequence complementary to the		
CC	complementary strand of a polynucleotide which comprises a 5'-end		
CC	sequence and an oligonucleotide comprising a sequence complementary to a		
CC	polynucleotide which comprises a 3'-end sequence, where the		
CC	oligonucleotide comprises at least 15 nucleotides and the combination of		
CC	the 5'-end sequence/3'-end sequence is selected from those defined in		
CC	the specification. The primer sets can be used in antisense therapy and		
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,		
CC	particularly full-length cDNAs. The primers are also useful for the		
CC	detection and/or diagnosis of the abnormality of the proteins encoded by		
CC	the full-length cDNAs. The primers allow obtaining of the full-length		
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and		
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to		
CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632		

XX 30-JUN-2000; 2000US-216503P.
 PR (MILL-) MILLENNIUM PHARM INC.
 XX Gluckemann MA;
 PI WPI; 2002-140091/16.
 XX N-PSDB; AAD27186.
 DR
 XX
 XX New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 XX Claim 4; Fig 1a; 103pp; English.
 PS
 XX The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as hematopoietic neoplastic disorders
 CC (e.g. leukemia), hematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology), and predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is human 57658 protein.
 CC
 XX
 SQ Sequence 277 AA;
 Alignment Scores:
 Pred. No.: 5,11e-272 Length: 277
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-896-522-3 (1-834) x AAE16592 (1-277)
 QY 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGAC 60
 DB 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGGGGCGCTTCCTGATAGGGGTGAGGGCGGCACTGCCAGCGGGAAGTGCACCTGTGT 120
 DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 QY 121 GAGAGATCATGAGTGTCTGGACAGAACAGAGGTGGAACAGCGGCGGAGAGTGGTC 180
 DB 41 GluValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnValVal 60
 QY 181 ATCTGAGCCAGGACAGGTTCTACAGAGTCTGACGCGACAGAGCAAGGCCAAGCCTTG 240
 DB 61 IleuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 241 AAAGAGACAGTCAATTTTGGACATCCAGATGCCCTTTGATATGATTTGATGACAGACT 300
 DB 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 QY 301 CTGAAGAACATCGGAGGCGCAAAACGTTGAGGTGCGACCTTATGATTTTGTGACACAC 360
 DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 QY 361 TCAAGTTACAGACGACCGGTGTACCTTCGCGACGTCGTCTGTGTGAGCGGATC 420
 DB 121 SerArgLeuProGluTThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

QY 421 TTGGTGTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGCGCCTTTCGTGACACC 480
 DB 141 LeuValPheTyrSerIleGlnGluIleArgAspMetPheHisLeuArgLeuPheValAlaPheThr 160
 QY 481 GACTCCAGAGTCAGGCTGTCTCGAAGAGTTCTCCGGAAGTGCAGCGCGGAGAGGACTTG 540
 DB 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 QY 541 GACCAATTTCTGACGCGATACACACCTTCGTGAAGCCGGCCTTCGAGAGATTTCCCTTG 600
 DB 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
 QY 601 CCGAGCAAGAGATGCCGATGTGATCATCCACGAGAGTGGACAAATGATGTCATC 660
 DB 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 QY 661 AACCTGATCGTGCAGACATCCAGACATTTCTGAATGTGACATCTGCAATGGCACCGA 720
 DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 QY 721 GAGAGGTCGAATGGCGGAGCTACAAAGCGACCTTTCTGAGCCAGGAGCAACCTGGG 780
 DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
 QY 781 ATGCTGACCTTCGGCAAGGTCACATTTGGAGTCCAGAGACACCCAC 831
 DB 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 RESULT 4
 ID AAM41288 standard; protein; 296 AA.
 AC AAM41288;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 6219.
 XX
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia.
 XX
 XX Homo sapiens.
 OS
 PN MO20015312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehtman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI60444.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -
 XX
 XX Example 2; SEQ ID NO 6219; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA438642-AA442213) with nootropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 XX Sequence 296 AA;

Alignment Scores:
 Pred. No.: 5,07e-272 Length: 296
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA441288 (1-296)

QY 1 ATGGCTTCGGCGGAGGCGAAAGCTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
 Db 20 MetAlaSerAlaGlyGlyAlaSerCysGluSerProAlaProGluAlaAspArgProHis 39
 QY 61 GACGGGCGCTTCCTGATGAGGGGGTGAAGGGGCGACGCGCGGAGGAGTGCACCGTCTC 120
 Db 40 GluArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 59
 QY 121 GAGAAGATCATGAGTGTCTGCGAGAGAAAGAGGTGAGAGCGGCGCGGAGAGTGTGTC 180
 Db 60 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 79
 QY 181 ATCTTGAGCGAGGACAGGCTTCTACAGGTCCTGACGCGAGAGGAGGCGGACCGCTTG 240
 Db 80 IleLeuSerGlnAspArgPheTyrLysValIleThrAlaGluGlnLysAlaLysAlaLeu 99
 QY 241 AAGGACAGTACAAATTTTGAACATCCAGATGCCCTTGATATGATTTGATGACAGAGACT 300
 Db 100 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnSpleuMetHisArgThr 119
 QY 301 CTGAAGACATCGTGGAGGCGCAAAACGCTGAGAGGTGCGGACCTATGATTTTGAACAC 360
 Db 120 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 139
 QY 361 TCAAGTTACAGAGACAGACGCGTGTCTACCTCGCGAGCGTGTCTGTTGAGGCGATC 420
 Db 140 SerArgLeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIle 159
 QY 421 TTGGTGTTCACAGCCGAGAGATCCGGGACATGTTCCACTGCGGCTCTTCTGAGACACC 480
 Db 160 LeuValPheTyrIleThrGlnGluIleArgAspMetPheHisIleuArgLeuPheValAspThr 179
 QY 481 GACTCCGACGTCAAGGCTGTCTCGAAGAGTCTCCGGGACGTGCGGAGGAGGAGACTG 540
 Db 180 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyArgAspLeu 199
 QY 541 GAGCAGATTCTGACGACGATACACCACTTCTGTAAGCGGCTTTCGAGAGATTCTGCTTG 600
 Db 200 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 219

QY 601 CCGCAAAAGAGTATCCGATGATCATCCAGAGAGTGAACATATGTTGGCATC 660
 Db 220 ProThrLysLysTyrLysAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 239
 QY 661 AACCTGATCGTGCAGACATCCAGACATTTGAAATGTGACATCTGCAATGGACCGA 720
 Db 240 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 259
 QY 721 GAGGGTCCAAATGAGGGGAGGAGCTTCAAGGCGGACTTTTTCGAGGCGAGGAGCCCTGGG 780
 Db 260 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerClnProGlyAspHisProGly 279
 QY 781 ATGCTGACCTTCGGCAAAAGCTGACATTTGAGTCCAGAGAGGAGCCGAC 831
 Db 280 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 296
 RESULT 5
 ID AA014412
 AC AA014412;
 DE 02-MAY-2002 (first entry)
 XX Protein of a human uridine kinase (UDK).
 DE
 XX
 XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001MO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI; 2001-626259/72.
 DR N-Psdb; AAK98735.
 XX
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance
 XX
 PS Claim 3; Page 29-30; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilising uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localisation studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and

CC immunological disorders. This sequence represents the protein of a human
XX uridine kinase of the invention.

XX Sequence 260 AA;

Alignment Scores:

Pred. No.:	9.2e-255	Length:	260
Score:	260.00	Matches:	260
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	93.86%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-3 (1-834) x AAO14412 (1-260)

```
QY 1 ATGGCTTCGGGGGAGGAGGAGAGTCCGAGAGCCCGCGGCGGAGCCGCTCCGAC 60
Db 1 MetAlaSerAlaGlyGlyGluAspGlySerProAlaProGluAlaAspArgProHis 20
QY 61 CAGCGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGGAAGTGCAGCGTGTGT 120
Db 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACAGCGGCGAGCGGAGGTGTC 180
Db 41 GlnLysIleMetGluLeuGlnGlnAsnGluValGlnGlnArgGlnArgLysValVal 60
QY 181 ATCTGGAGCCGAGCAGGTTTACAGGCTCTGACGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 241 AAAGACAGTCAATTTTGAACATCCAGATGCTTTGATATGATTTGATGCACAGACT 300
Db 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 301 CTGAAGAACATCGTGGAGGGGCAAAACGGTGGAGGTGCCGACCTATGATTTTGACACAC 360
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 361 TCAGAGTTACCGAGACACCGGTGTCTACCCCTCGGACCGTGTCTGTTGAGGGCATC 420
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGATGCACACC 480
Db 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuAlaGlyLeuPheValAspThr 160
QY 481 GACTCCGACGTGAGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGGAGACTG 540
Db 161 AspSerAspValAlaArgLeuSerArgValLeuAlaArgAspValAlaArgGlyArgAspLeu 180
QY 541 GAGCAGATTCTGACGAGTACACACCTTCGTGAAGCGCGGCTTCGAGAGTTTGCTG 600
Db 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY 601 CCGCAAGAAGTATGCGCATGTATCATCCACGAGAGTGAACAATATAGTTCATC 660
Db 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
QY 661 AACCTGATCGTGACGACATCCAGAGCAATTCGATAGTGAATGTCGAATGGCACCGA 720
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
QY 721 GGAGGGTCCCATGGGCGGAGCTACAAGCGACCTTTCTGAGCCAGGAGGACCAACCTGGG 780
Db 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGlnProGlyAspHisProGly 260
```

RESULT 6

AA664506 standard; Protein, 276 AA.

XX AA664506;
XX AC
XX

DT 02-OCT-2001 (first entry)

XX Human uridine kinase.

XX Human; uridine kinase; UK.

XX Homo sapiens.

XX CN1287172-A.

XX 14-MAR-2001.

XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.

XX (UYFU-) UNIV FUDAN.

XX Yu L, Zhao Y, Zhang H;

XX WPI: 2001-409529/44.

XX N-PSDB; AAH75355.

XX Human uridine kinase and its coding sequence, preparation and

XX application -

XX Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.

XX The invention relates to human uridine kinase (UK).

XX Sequence 276 AA;

Alignment Scores:	1.04e-182	Length:	276
Pred. No.:	189.00	Matches:	189
Score:	189.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	68.23%	Gaps:	0

US-09-896-522-3 (1-834) x AA664506 (1-276)

```
QY 172 AAGGTGTCATCTGAGCCAGAGACAGGTTCTACAGGTCCTGACGCGACAGAGGCC 231
Db 57 LysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnLysAla 76
QY 232 AAGGCTTGAAAGACAGTCAATTTTGAACATCCAGATGCGCTTGATATGATTTGATG 291
Db 77 LysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMet 96
QY 292 CACAGAGCTCTGAAGAACATCGTGGAGGGCAAAACGGTGGAGGTGCGACCTATGATTTT 351
Db 97 HisArgThrLeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPhe 116
QY 352 GTGACACACTCAAGGTTTACAGAGACACAGCGTGGTCTACCTTGGGACGTGTTCTGTTT 411
Db 117 ValThrHisSerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPhe 136
QY 412 GAGGGCATCTTGTTGTTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTCTTC 471
Db 137 GlnGlyIleLeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPhe 156
QY 472 GTGAGACCGGACTCGAGCGTCAAGGCTGTCCGAGAGTTCTCCGGGACGTGCGCGAGGG 531
Db 157 ValAspThrAspSerAspValAlaArgLeuSerArgValLeuArgAspValAlaArgGly 176
QY 532 AGGACCTGAGAGCATTTCTGACGACGATACACCACTTGTGAAAGCGGCTTCGAGAG 591
Db 177 ArgAspLeuGlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGln 196
QY 592 TTCTGCTTGGCCGCAAAAGATGTGCGATGTGATATCCACGAGAGTGAACAATATG 651
Db 197 PheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspAsnMet 216
```

QY 652 GTTGCATCAACCTGATCGTGAGACATCCAGAGATTGATGTCATCTGCAAA 711
 |||
 Db 217 ValAlaIleAsnLeuIleValGlnHisIleGlnHisIleLeuValAsnGlyAspIleCysLys 236
 |||
 QY 712 TGGCACCAGAGAGGTCCTCAATGGGCGG 738
 |||
 Db 237 TrpHisArgGlyGlySerAsnGlyArg 245
 |||
 RESULT 7
 ABB89353
 ID ABB89353 standard; Protein; 190 AA.
 XX
 AC ABB89353;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1729.
 XX
 XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;
 KM antileptergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
 KM vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 XX 19-MAY-2000; 2000US-205515P.
 ER
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Birse CE, Rosen CA;
 XX
 DR WPI, 2002-122018/16.
 DR N-PSDB; ABL89762.
 XX
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive, and
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 XX Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
 PS
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WFO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 190 AA;

Alignment Scores: 6.62e-129 Length: 190
 Pred. No.: 136.00 Matches: 136
 Score:

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 49.10% Indels: 0
 DB: 23 Gaps: 0
 US-09-896-522-3 (1-834) x ABB89353 (1-190)
 QY 100 AGCGGAGAGTCGACCGTGTCGAGAGATCATGAGATTGCTGGACAGAGAGAGTGGA 159
 |||
 Db 34 SerGlyLysSerThrValCysGluLysIleMetGlnLeuLeuGlnAsnGlnValGln 53
 |||
 QY 160 CAGCGGACAGCGAGAGGTGTCATCTGAGCCAGAGACAGGTTCTACAGGTCCTGACGGCA 219
 |||
 Db 54 GlnArgGlnArgLysValValIleLeuSerGlnAspArgPheTyrLysValIleThrAla 73
 |||
 QY 220 GAGCAGAGAGCCAGGCTTTGAAAGACAGTCATATTTTGACCATTCAGATGCTTTGAT 279
 |||
 Db 74 GluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAsp 93
 |||
 QY 280 AATGATTTGATGACAGCACTGTAAGAACATCTGAGAGGCAAAAGGTCGAGGTCGCG 339
 |||
 Db 94 AsnAspLeuMetHisArgThrLeuLysAsnIleValGlnGlyThrValGlnValPro 113
 |||
 QY 340 ACCTATGATTTTGTGACACACTCAAGTTACAGAGACAGACAGGTCGCTTACCTGCGGAC 399
 |||
 Db 114 ThrTyrAspPheValThrHisSerArgLeuProGlnThrThrValValTyrProAlaAsp 133
 |||
 QY 400 GTGGTTCTGTTTGGAGGCACTTGTGTTCTACAGCCAGAGAGATCCGGACATGTTTCCAC 459
 |||
 Db 134 ValValLeuPheGluGlnGlyIleLeuValPheTyrSerGlnGlnIleArgAspMetPheHis 153
 |||
 QY 460 CCGCGGCTCTGTGTGACACCGACTCCGACGTCGAGCTGTCGAGGA 507
 |||
 Db 154 LeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
 |||
 RESULT 8
 ABB89328
 ID ABB89328 standard; Protein; 120 AA.
 XX
 AC ABB89328;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human MDDT protein Incyte ID No. LI:235557.12.orf2:2001JAN12.
 XX
 KM Human; molecule for disease detection and treatment; MDDT; cancer;
 KM cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KM autoimmune disorder; inflammatory disorder; Crohn's disease;
 KM multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KM hepatotropic; immunosuppressive; antiasthmatic.
 XX
 OS Homo sapiens.
 XX
 PN WO200255738-A2.
 XX
 PD 18-UTL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US01008.
 XX
 PR 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261655P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262336P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman UT, Jones AL;
 PI Dam TC, Liu TF, Harris B, Flores V, Dafio A, Marwaha R, Chen AJ;
 PI Chang SC, Geretin EH, Peralta CH, David MH, Lewis SA;

XX MPI: 2002-590679/63.
DR N-FSDB; ABS51801.

XX
PT New disease detection and treatment molecule (MDDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDDT expression, e.g. autoimmune or
PT inflammatory disorders -

PS Claim 27, Page 120, 129pp; English.

XX
CC The present invention relates to the isolation of novel human
CC molecules for disease detection and treatment (MDDT), and the
CC polynucleotide sequences (mddt) encoding them. The MDDT polypeptides
CC may be used to screen for molecules that bind to, or are bound by the
CC encoded polypeptides, and to develop a transcript image of a tissue or
CC cell type. Probes comprising at least 20 nucleotides of the mddt
CC polynucleotide may be used to assess the toxicity of a test compound.
CC The MDDT polypeptides and mddt polynucleotides are useful in the
CC diagnosis, study, prevention and treatment of diseases associated with
CC the expression of molecules for disease detection and treatment. Such
CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC cirrhosis, or cancers) and autoimmune/inflammatory disorders
CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mddt
CC polynucleotides may also be used as molecule markers, in microarrays,
CC and in somatic or germ-line gene therapy. ABG70306-ABG70341 represent
CC the MDDT proteins of the invention.

5Q	Sequence	120 AA;
Alignment Scores:		
Pred. No.:	1.59e-104	Length: 120
Score:	112.00	Matches: 112
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	40.43%	Indels: 0
DB:	23	Gaps: 0

US-09-896-522-3 (1-834) x ABG70328 (1-120)

QY	268	GATGCCCTTATATATGATTTGATGACACAGACTTGAAAGACATGCTGAGGCGAAAACG	32
Db	5	AspIaIaheAspAnsbuLeuWethIaSaGThrLeuVAsnIleValGIuGIlySerThr	24
QY	328	GTGAGAGGTGCACCACTATGATTTTGGACACACTCAAGGTTCAGAGACCAGGTGTC	38
Db	25	ValGIuValIProThrIlyrAspPheValIInrHisSerArgLeuProGIuInrThValVal	44
QY	388	TACCTTCGGACGCGNGTCTTGTTTGGAGGCATCTTGTTCTTAACGCAGAGATCCGG	44
Db	45	TyrProIaAspValValLeuPheGIuGIlyIleLeuValPheTyrSerGIuGIuIleArg	64
QY	448	GACATGTTCCACCTCGCGCCTCTTCGTGACACCGCATCCGACGTGACGAGCTGTCTCGAAGA	50
Db	65	AspPheCpheiIleuArgLeuPheValIaPThrAspSerAspValArgLeuSerArgArg	84
QY	508	GTTTCTCCGGGACGTGGCCGAGGAGGAGACCTGAGACAGATTCTGACCGAGTAACAAC	56
Db	85	ValLeuArgAspValArgArgGIyArgAspLeuGIuGIuIleLeuThrGIuGIuIlyThr	104
QY	568	TTGCGTAAGCCGGGCTTCGAGGAGTTCGCTCCGCCG	603
Db	105	PheValIySerProIaPheGIuGIuPheCysLeuPro	116

RESULT 9	
AA014413	
ID	AA014413 strand; Protein; 277 AA.
XX	
XX	
AC	AA014413;
DT	
02-MAY-2002	(first entry)
XX	
XX	
XX	Protein relating to a human uridine kinase (UDK) of the invention

XX	Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KM	probe; chromosome localisation study; tissue expression; gene therapy;
KM	antibody; vaccine; human ovarian cancer; immunological disorder;
KW	human colon carcinoma; immunogen.
XX	
OS	unidentified.
XX	
PN	WO200172963-A2.
PN	
PD	04-OCT-2001.
XX	
XX	
PF	27-MAR-2001; 2001WO-US09663.
XX	
PR	27-MAR-2000; 2000US-0536647.
XX	
PA	(SMIK) SMITHKLINE BEECHAM CORP.
XX	
PI	Ho YS, Johnson RK;
XX	
DR	WPI; 2001-626259/72.
XX	
PT	Novel human uridine kinase polypeptides useful for treating cancers,
PT	and to identify agonists and antagonists of the polypeptide useful for
PT	treating conditions associated with uridine kinase imbalance
XX	-
XX	
PS	Disclosure; Page 23; 31pp; English.

CC The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilizing uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC NOTE: The present sequence is stated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.

SO	Sequence	277	AA:
Alignment Scores:			
Pred.	No.:	1,42e-36	length:
Score:	No.:	45.00	Matches:
Percent Similarity:		100.00%	Mismatches:
Best Local Similarity:		100.00%	Indels:
Query Match:		16.25%	Gaps:
DB:		22	

[illegible]

Qy 418 ATCTGCTGTCTAC 432
 |||||
 Db 140 IleuValPheTyr 144

RESULT 10
 AAM38694
 ID AAM38694 standard; Protein; 261 AA.

AC AAM38694;
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1839.
 XX
 KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 KW
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0634450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AAI57850.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 3; SEQ ID NO 1839; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 SQ Sequence 261 AA;

Alignment Scores: 2.84e-16 Length: 261
 Pred. No.:

Score: 25.00 Matches: 25
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 9.03% Mismatches: 0
 DB: 22 Indels: 0
 Gaps: 0

US-09-896-522-3 (1-834) x AAM38694 (1-261)

Qy 565 ACCTTGCTGAAGCGGCTTCGAGGAGTCTGCTGCCGACAAAGATATGCCGATGTG 624
 |||||
 Db 187 ThrPheValIysProAlaPheGluGluPheCysLeuProThrIlysrAlaApVal 206

Qy 625 ATCATCCACGAGGA 639
 |||||
 Db 207 IleIleProArgGly 211

RESULT 11
 AAB73494
 ID AAB73494 standard; Protein; 261 AA.
 AC AAB73494;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Human transferase HTFS-1, SEQ ID NO:1.
 XX
 KW Human transferase; HTFS; agonist; antagonist; cellular signalling;
 KW proliferation; cell proliferative disorder; immune disorder;
 KW atherosclerosis; hepatitis; psoriasis; cancer; tumor;
 KW inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
 KW cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
 KW multiple sclerosis; rheumatoid arthritis; pancreatitis;
 KW systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
 KW haemodialysis; extracorporeal circulation; trauma; transgenic animal;
 KW gene therapy; drug screening.
 KW
 OS Homo sapiens.
 XX
 PN WO200132888-A2.
 XX
 PD 10-MAY-2001.
 XX
 PF 02-NOV-2000; 2000MO-US30485.
 XX
 PR 04-NOV-1999; 99US-0163595.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
 PI Shih JL, Azimzai Y, Lu DM, Baughm MR;
 XX
 DR WPI; 2001-328796/34.
 DR N-PSDB; AAM23801.
 XX
 PT Human transferase polypeptides and polynucleotides useful for
 PT diagnosis, prevention and treatment of cell proliferative and immune
 PT system disorders and for identifying agonists and antagonists -
 XX
 PS Claim 1; Page 103-104; 157pp; English.
 XX
 CC Sequences AAB73494-AAB73535 represent novel human transferase proteins
 CC HTFS-1 to HTFS-42, and sequences AAM23801-AAM23842 represent cDNAs
 CC encoding them. The proteins play important roles in the regulation of
 CC cellular signalling and proliferation. The HTFS proteins are useful for
 CC screening compounds for their effectiveness as agonists or antagonists of
 CC transferase activity, or for compounds that specifically bind to an HTFS
 CC protein or which modulates the activity of an HTFS protein.
 CC Pharmaceutical compositions comprising an HTFS protein, HTFS
 CC agonist or antagonist, or genetic construct encoding an HTFS
 CC protein are useful for treating a disease or condition associated
 CC with decreased or increased expression of functional HTFS. Disorders
 CC which may be treated using such compositions include cell proliferative
 CC disorders and immune disorders. For example, diseases which may be

CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HSPs proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HSPs, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HSPs may be
CC used for the diagnosis of disorders associated with the expression of
CC HSPs, or in assays to monitor patients being treated with HSPs or
CC agonists, antagonists or inhibitors of HSPs. The present sequence
CC represents an HSPs protein of the invention.

XX Sequence 261 AA;

Alignment Scores:

Pred. No.:	2,846-16	Length:	261
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-3 (1-834) x AAB73494 (1-261)

QY 565 ACCTTGGAAGCCGCGCTTCGAGAGTTCTGCTGCCGCAAGAAGATATGCCATGTG 624

Db 187 ThPheVallylsProAlaPheGluGluPheCysLeuProThrIysTyAlaAspVal 206

QY 625 ATCATCCACGAGGA 639

Db 207 IleIleProArgIly 211

RESULT 12

AAB56582 standard; Protein: 337 AA.

XX AAB56582;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.

XX Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

DR WPI, 2000-587513/55.

DR N-PSDB; AAF15785.

XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

XX Claim 11; Page 1566-1567; 2338pp; English.

XX AAF1566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56163 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotoxic, antineoplastic, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 337 AA;

Pred. No.:	2,766-16	Length:	337
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	21	Gaps:	0

US-09-896-522-3 (1-834) x AAB56582 (1-337)

QY 565 ACCTTGGAAGCCGCGCTTCGAGAGTTCTGCTGCCGCAAGAAGATATGCCATGTG 624

Db 263 ThPheVallylsProAlaPheGluGluPheCysLeuProThrIysTyAlaAspVal 282

QY 625 ATCATCCACGAGGA 639

Db 283 IleIleProArgIly 287

RESULT 13

ABP41393 standard; Protein: 337 AA.

XX ABP41393;

AC ABP41393;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOBLP29, SEQ ID NO:2525.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 1p32.

XX Homo sapiens.

PN WO200200677-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001WO-US18569.

PR 07-JUN-2000; 2000US-209467P.

PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 XX N-PSDB; ABO54470.
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID No 2525; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovarian and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 337 AA;
 Alignment Scores:
 Pred. No.: 2,766-16 Length: 337
 Score: 25.00 Matches: 25
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.03% Indels: 0
 DB: 23 Gaps: 0
 US-09-896-522-3 (1-834) x ABP41393 (1-337)
 QY 565 ACCTTCGTGAAGCGGCTTCGAGAGATTCTGCTGCCGACAAAGATATGCCGATGG 624
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 263 ThpHeVallyspRoAlaPhgUgluPhcYsleuProThrlyslsTYAlaAspVal 282
 QY 625 ATCATCCACGAGGA 639
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 283 lIeIlleProArgGly 287
 RESULT 14
 AAB16594
 ID AAB16594 standard; Protein; 125 AA.
 XX
 AC AAB16594;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 protein prodomain consensus sequence #1.

XX
 KW Human: uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteoporotic; cardiac; neuroprotective; nootropic;
 KW anticonvulsant; prodomain.
 XX
 OS Homo sapiens.
 XX
 PN WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 DR WPI; 2002-140091/18.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 PS Disclosure; Fig 4; 103pp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays).
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is the consensus sequence of the biological sample.
 CC prodomain, kinase uridine monophosphokinase transferase ATP-binding
 CC kinase-like ribonucleoside pyrimidine FIS cDNA domain.
 CC
 XX
 SQ Sequence 125 AA;
 Alignment Scores:
 Pred. No.: 4,356-06 Length: 125
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.42% Indels: 0
 DB: 23 Gaps: 0
 US-09-896-522-3 (1-834) x AAB16594 (1-125)
 QY 736 CGAGCTACAGCGGACCTTTTCTGAGCGAGGAGACACCTGGG 780
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 94 ArgSerTyllyArgThrpheSerGluProGlyAAspHisProGly 108
 RESULT 15
 ABB62307
 ID ABB62307 standard; Protein; 260 AA.
 XX
 AC ABB62307;
 XX

```

XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster polypeptide SEQ ID NO 13713.
XX DE
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL06410.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI6175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 260 AA;

Alignment Scores:
Pred. No.: 4e-06 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.42% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x ABB62307 (1-260)

OY 388 TACCTGCGGACGAGTCTGTTGAGGACATCTGGTGTCTAC 432
Db 135 TyrProAlaAspValValLeuPheGluCylLeuValPheTyr 149

```

Search completed: November 25, 2003, 07:53:19
 Job time : 34.5012 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 06:52:49 ; Search time 48.1806 Seconds
(without alignments)
8933.708 Million cell updates/sec

Title: US-09-896-522-3

Perfect score: 1514
Sequence: 1 atgcgtctgcgcggagcgca.....ccagcagcagaccactga 834

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 161050

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO/US09896522/runat_21112003_184104_2821/app_query.fasta_1.2830
-DB=SPTREMBL_23 -QFMT=fasta -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.ccn_1_1.182.0/runat_21112003_184104_2821 -NCPV=6 -ICPV=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELACK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertedrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020.5	67.4	201	4 Q96BJ0	Q96bj0 homo sapien

2	932	61.6	261	11 Q8C476	Q8c476 mus musculi
3	733	48.4	260	5 Q8MRJ1	Q8mrj1 drosophila
4	570.5	37.7	486	10 Q9KFS0	Q9kfs0 arabidopsis
5	568	37.5	465	10 Q9LTV6	Q9lrv6 arabidopsis
6	563	37.2	483	10 Q9LTK3	Q9ltk3 arabidopsis
7	559	36.9	515	5 Q19583	Q19583 caenorhabdi
8	559	36.9	555	5 Q9U317	Q9u317 caenorhabdi
9	544	35.9	466	10 Q8VYB2	Q8vyb2 arabidopsis
10	542	35.8	614	5 Q9V810	Q9v810 drosophila
11	542	35.8	626	5 Q8MOK4	Q8mok4 drosophila
12	540.5	35.7	419	4 Q8N524	Q8n524 homo sapien
13	529	34.9	469	10 Q8LD95	Q8ld95 arabidopsis
14	507.5	33.5	542	10 Q9LQ32	Q9lq32 arabidopsis
15	506	33.4	1060	10 Q9LFFZ2	Q9lffz2 arabidopsis
16	459	30.3	209	16 Q8E0A8	Q8e0a8 streptococc
17	459	30.3	454	3 Q74427	Q74427 schizosacch
18	458	30.3	209	16 Q8E5Y5	Q8e5y5 streptococc
19	451	29.8	207	16 Q8CSB2	Q8csb2 staphylococ
20	441	29.1	209	16 Q8DTG1	Q8dtg1 streptococc
21	439	29.0	419	5 Q9BMX4	Q9bmux4 cryptospori
22	404	26.7	210	16 Q8EPT5	Q8ept5 oceanobacil
23	403	26.6	213	16 Q8BD84	Q8bd84 vibrio vuln
24	390.5	25.8	111	4 Q9BU42	Q9bu42 homo sapien
25	380.5	25.1	111	4 Q92528	Q92528 homo sapien
26	369	24.4	105	11 Q9OYGH	Q9oygh rattus norv
27	349	23.1	212	16 Q8EDY4	Q8edy4 shewanella
28	233	15.4	674	10 Q9C664	Q9c664 arabidopsis
29	212	14.0	643	10 Q9C9B9	Q9c9b9 arabidopsis
30	210.5	13.9	304	16 Q8YPR9	Q8ypr9 anabaena sp
31	210.5	13.9	403	16 Q8GRU9	Q8gru9 oryza sativ
32	207	13.7	646	10 Q8S3R9	Q8s3r9 oryza sativ
33	206.5	13.6	366	10 Q9S033	Q9s033 odontella s
34	200	13.2	334	2 Q8L0S5	Q8l0s5 synchococc
35	200	13.2	334	16 Q8DHN2	Q8dhn2 synchococc
36	199.5	13.2	352	10 P93681	P93681 pium sativ
37	197	13.0	367	3 Q94642	Q94642 schizosacch
38	196.5	13.0	448	10 Q8GUE1	Q8gue1 galidieria s
39	195.5	12.9	333	2 Q9LBU7	Q9lbu7 synchococc
40	191	12.6	313	16 Q8YUJ6	Q8yuj6 anabaena sp
41	187.5	12.4	271	10 Q9SPH7	Q9sph7 beta vulgar
42	178.5	11.8	232	3 Q12084	Q12084 saccharomyc
43	172	11.4	555	16 Q9WZL0	Q9wzl0 thermotoga
44	169.5	11.2	405	10 Q9ATC3	Q9atc3 vaucheria 1
45	157.5	10.2	542	2 Q68872	Q68872 myxococcus

ALIGNMENTS

RESULT 1

Q96BJ0 PRELIMINARY; PRT; 201 AA.

AC Q96BJ0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL EMBL; BC015547; AAH15547.1; -
DR InterPro: IPR006083; PRK URK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK, 1.
DR PRINTS: PR00988; URIDINKINASE.
KW KINASE.

SEQUENCE 201 AA; 22790 MW; 0E5FF2F00FF7B5363 CRC64;

Alignment Scores:

Pred. No.: 2,586-80 Length: 201
 Score: 1020.50 Matches: 200
 Percent Similarity: 85.47% Conservative: 0
 Best Local Similarity: 85.47% Mismatches: 1
 Query Match: 67.40% Indels: 33
 DB: 4 Gaps: 1

US-09-896-522-3 (1-834) x Q96BU0 (1-201)

QY 1 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
 DB 1 MetlaeserlaaglygluapCysgluSerProlaaproglyalaapargProHis 20
 QY 61 CAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCGCAGCGGGAAGTGCACCTGTGT 120
 DB 21 GlnarproPheleuilegilyalsercllygilythrilaSerglylserThrValCys 40
 QY 121 GAGAAATCATGAGTGTCTGGAAGAGAGGTTGGAACAGCGGCGGAGCGGAGGAGTGTG 180
 DB 41 GlnulylleuSerGluLeuLeuGlyGlnabnglylValGlnGlnarGlnarGlylVal 60
 QY 181 ATCTGAGCCAGAGCAGGTTTCAAGAGTCTGACGCGAGAGCAAGCGCCAGCGCTTG 240
 DB 61 IleuSerGlnaapargPheTyrlylValleuThrilaGlnGlnlylValalaValleu 80
 QY 241 AAAGACAGTCAATTTTACCATTCAGATCCCTTTGATTAATGATTTGATGACAGAGCT 300
 DB 81 LysGlyGlnlyrAanPheaspHisProaspalaPheaspAaspLeuMetHisargThr 100
 QY 301 CTGAAGAACATGTCGAGGCGGCAAAACGCTGAGGCGGAGCGGACCTTATGATTTTGTGAC 360
 DB 101 LeuylsbenilevalGlnGlylylserThrValGlnValProthlyrAaspPheValThHis 120
 QY 361 TCAAGTTACAGAGACACCGGCGTGTCTACCTCGGACGCTGTTCTTTGAGGCGATC 420
 DB 121 SerarGlnuProGlnuThrThrValValTyrProalaapValValleuPheGlnGlyle 140
 QY 421 TTGGTGTCTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGGCTCTTGTGTGACACC 480
 DB 141 LeuValPheTyrserGlnGlnlylAargaspMetPheHisleuAargLeuPheValAaspThr 160
 QY 481 GACTCCGAGCTGAGGCGTGTCTCGAAGAGTTCTCGGAGCGTGGCCGAGGAGGAGCCTG 540
 DB 161 AaspseraspValaargLeuSerArg--Arg----- 169
 QY 541 GAGCAGATTCTGACGAGTACACCACTTCGTGAAGCGGCTTCGAGAGTTCTGCTG 600
 DB 169 ----- 169
 QY 601 CCGAAGAAAGATATGCCGATGTGATCATCCCAAGAGAGTGGACATATGTGTCATC 660
 DB 170 --AaspLysGlnValCysargCysaspHisProthlyrAargSerGlyGlnTyrclYcHis 189
 QY 661 AACCTGATCGTGCAGCAGCATCCAGACATCTTGATGAG 698
 DB 189 InProaspargAlaAlaHisProGlyHisSerGlnuTyr 201

RESULT 2
 Q8C476 PRELIMINARY; PRT; 261 AA.
 AC Q8C476;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Uridine-cytidine kinase 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;

RX MEDLINE=22354683; PubMed=12466851;
 RA The RANOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK082837; BAC38646.1; "-
 SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA0153 CRC64;

Alignment Scores:

Pred. No.: 1,416-72 Length: 261
 Score: 932.00 Matches: 187
 Percent Similarity: 79.04% Conservative: 28
 Best Local Similarity: 68.75% Mismatches: 41
 Query Match: 61.56% Indels: 16
 DB: 11 Gaps: 4

US-09-896-522-3 (1-834) x Q8C476 (1-261)

QY 22 GACTGGAGAGCCCGCGCGGAGCGGAGCGGACCGTCCGAC---CAGCGGCTTCTGATA 78
 DB 4 AaspserGlnGlnThrleuGlnabnglylGlnGlnProaspGlyGlyGluProPheleuile 23
 QY 79 GGGGTAGCGCGGCGGACCTGCGAGCGGAGTGCACCTGTGTGAGAAATCATGAGAGTTG 138
 DB 24 GlyValsercllygilythrilaSerGlylserSerValCysAlaValleuGlnleu 43
 QY 139 CTGGACAGAACAGAGTGAACAGCGGCGGAGAGTGTCTATCTTGACCGACAGCAG 198
 DB 44 LeuGlyGlnabnglylValaaspTyrHisGlnlylGlnValleuSerGlnaPheSer 63
 QY 199 TTCTACAGGCTCTGAGCGGAGCGGAGGCGGAGCGGACCTTGAAGACATGATTTT 258
 DB 64 PheTyrargValleuThrserGlnGlnlylValalaPheleuSerGlylPheleuSer 83
 QY 259 GACCATCAGATGCTTTGATTAATGATTTGATGACAGACTCTGAAGACATCTGTGAG 318
 DB 84 AspHisProaspalaPheaspAaspGlnleuilePheValThrleuSerGlnuThrGln 103
 QY 319 GGCAGAAACGCTGAGGTCGCCGACCTTGTGATTTTGTGACACACTCAAGATTACAGAGACC 378
 DB 104 GlylserThrValGlnleuProValTyrAaspPheValSerHisSerargLysGlnuThr 123
 QY 379 ACGTGTGTACCTCGGAGCGGAGTGTCTGATTTGAGGAGTGTGTCTTCAAGCGCAG 438
 DB 124 ValThrleuTyrProalaapValValleuPheGlnGlyleuValaPheTyrserGln 143
 QY 439 GAGATCCGGAGACATGTTCCACCTGCGGCTTCTGTGAGACAGCACTCCGACGTCAAGCTG 498
 DB 144 GlnValAargaspLeuPheGlnleuTyrleuPheValAaspThrAaspThrAargLeu 163
 QY 499 TCTGAAAGATTCTCGGAGCGT---CGCCGAGGAGGAGCCTGAGACAGATTCTGACG 555
 DB 164 SerargargValleuAargaspHisSerGlnargGlyAargAaspLeuGlnuThrleuSer 183
 QY 556 CAGTACACACCTTGTGAAGCGGCTTCGAGGATTTGCTGCGCGGACAAAGAGTAT 615
 DB 184 GlnTyrThrPheValLysProalaPheGlnuThrPheCysleuProthlyrlySerlyr 203
 QY 616 GCCGATGTATCATCCACAGAGAGTGAACAATATGTTGCCATCAACTGATCTGACG 675
 DB 204 AlaaspValleuThrleuProarglylalaPhePheleuValAlaileuSerleuileValGln 223
 QY 676 CACATCCAGACATTTGTAATGTGTGACATCTGCAAAATGGACCGAGAGAGTCCAAATGG 735
 DB 224 HisleuGlnaspilleuasnngly-----Gly 232
 QY 736 CGAGGTACAAGCGGACCTTTCTGAGCAGGAGGACCACTGGAGTGTGACTGTGGC 795
 DB 233 LeuSerlyserGlnThr-----AasnGlyTyrPheasnGlyTyrThrProSerArg 249
 QY 796 AACGGTCACTTTGAGTTCAGACAGACAGCCAC 831

Db 250 LysArgGlnAlaSerGlnSerSerSerArgProHis 261

RESULT 3

Q8MRJ1 PRELIMINARY; PRT; 260 AA.

AC Q8MRJ1; 260 AA.

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE L013909P.

GN C6364.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,

RA Champ M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez W., Gharin H., Krommiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Ceiniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY119583; AAM50237.1; -.

DR FlyBase; FBgn0039179; CG6364.

DR InterPro; IPR006083; PRK_URK.

DR InterPro; IPR007664; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PR00988; URIDINKINASE.

DR SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A1655 CRC64;

Alignment Scores:

Pred. No.: 3.03e-55 Length: 260

Score: 733.00 Matches: 142

Percent Similarity: 82.46% Conservative: 32

Best Local Similarity: 67.30% Mismatches: 35

Query Match: 48.41% Indels: 2

DB: Gaps: 2

US-09-896-522-3 (1-834) x Q8MRJ1 (1-260)

QY 67 CCCTCTCTGTAAGGGGTGAGCGCGGCACTGCCAGCGGAACTGCTGTGAGAG 126

Db 27 ProPheLeuIleGlyValAlaGlyThrAlaSerGlySerThrValCysLysLys 46

QY 127 ATCATGAGTGTGCGGAGAGAAAGAGTGAACAGCGGAGCGGAGTGGCATCTG 186

Db 47 IleMetGlnGlnLeuGlyGlnAlaGlnMetCaspHisThrGlnMglnValIleSerIle 66

QY 187 AGCCAGGAGAGGTTCTTCAAGAGTCTTACGCGAGAGAGAGGCAAGGCTTGAAGA 246

Db 67 SerGlnAspSerPheTyrArgGlnLeuThrProAlaGlnLysAlaLysAlaGlnLysGly 86

QY 247 CAGTACATTTTGAACCATCCAGATGCTTGTATATATTGATGACAGAGACTTGAAG 306

Db 87 LeuPheAsnIlePheAspHisProAspAlaPheAsnGlnLysLeuMetTyrSerThrLeuGln 106

QY 307 AACATCGTGGAGGCAAAACGTTGAGAGTCCGACCTATGATTTTGACACACTCA--- 363

Db 107 AsnIleLeuLysGlyHisLysValGlnIleProSerTyrAspTyrArgThrAsnSerLeu 126

QY 364 AGGTTACAGAGACACAGCGTGTCTTACCTGCGAGCGTGTCTTGTGAGGCATCTTG 423

Db 127 AspPheGlnAsnValIleLeuValIleTyrProAlaAspValIleLeuPheGlnGlyIleLeu 146

QY 424 GTGTTCTACAGCCAGGAGATCCCGGACATGTTCCACTGCGCTCTTCTCGGACACCGAC 483

Db 147 ValPheTyrPheProLysIleArgGlnLeuPheIleMetLysLeuPheValAspThrAsp 166

QY 484 TCCGACGTACAGGCTGTCTCGAAGAGTTCTCCGGAGCGTG---CGCCGAGGAGGAGCGTG 540

Db 167 ProAspThrArgLeuAlaArgValProArgAspIleAsnGlnLysGlyArgAspLeu 186

QY 541 GAGCAGATTCTGACGAGATACACACCTTCGTGAAGCGGCTTCGAGAGATTCGCTG 600

Db 187 AspAlaValLeuThrGlnTyrMetThrPheValIleProAlaPheGlnGlnPheCysSer 206

QY 601 CCGACAAAGAGAGTACCGCATGATCCATCCGACGAGAGTGAACAATAGTTCGATC 660

Db 207 ProThrLysPheAsnValIleLeuProArgGlyAlaAspAsnThrValAlaIle 226

QY 661 AACCTATCGTGCAGACATCCAGACATCTTG 693

Db 227 AspLeuIleValHisHisIleGlyGlnIleLeu 237

RESULT 4

Q9FKS0 PRELIMINARY; PRT; 486 AA.

AC Q9FKS0;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Uridine Kinase-like protein.

GN AT5G40870.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Euroside II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RA MEDLINE=98344145; PubMed=9679202;

RA Kaneo T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,

RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence

RT features of the regions of 1,381,565 bp covered by twenty one

RL physically assigned p1 and TAC clones.";

RL DNA Res. 5:131-145(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan Y.W., Lee J.M.,

RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,

RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,

RA Palm C.U., Shim P., Southwick A., Tripp M.G., Wu T., Davis R.W.,

RA Ecker J.R., Theologis A.;

RT "Arabidopsis Open Reading Frame (ORF) Clones.";

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB011477; BAB11349.1; -.

DR EMBL: BT002336; AAN86169.1; -.

DR HSSP; Q26998; 1BD3.

DR InterPro; IPR006083; PRK_URK.

DR InterPro; IPR00836; PRTtransferase.

DR InterPro; IPR00764; Uridine_kin.

DR Pfam; PF00156; Priboyltran_1.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PR00988; URIDINKINASE.

DR TIGRPFAM; TIGR00235; udk; 1.

DR Kinase; Transferase.

SO SEQUENCE 486 AA; 54430 MW; 50DA1CB89346FB54 CRC64;

Alignment Scores:

Pred. No.: 5.09e-41 Length: 486

Score: 570.50 Matches: 119

Percent Similarity: 66.40% Conservative: 49

Best Local Similarity: 47.04% Mismatches: 68

Query Match: 37.68% Indels: 17

DB: Gaps: 7

US-09-896-522-3 (1-834) x Q9FKS0 (1-486)

QY 31 AGCCCGCGCCGAGGCGAGCGTCCGAC-----CAGCGGCTTCTGTATAGGGGTG 84

```

Db      48 SerSerSerProSerSerSerAspProGluAlaProIlysglnProPheIleIleGlyVal 67
Qy      85 AGCGCGGCGACTGCCAGCGGGAAGTCCAGCCTGTGTGAGAAAGATCATGAGACTTCTGGA 144
Db      68 SerIlyGlyThrIleAspSerGlyLysThrThrValCysAspMetIleIleGlnGlnLeu--- 86
Qy      145 CAGAAACGAGGTGAAACAGCGGACCGGAGGAGTGTGATCCTCGAGACCGAGACAGGTGTAC 204
Db      87 -----HisAspHisArgValAlaLeuValAlaAsnGlnAspSerPheTyr 100
Qy      205 AAGGCTCTGACGAGACAGAGACAGAGCGCCCTTGAAAGACAGTCAATTTTGACCAT 264
Db      101 ArgGlyLeuThrSerGlnGlnLeuGlnArgVal-----GlnIlyTyrAsnPheAspHis 118
Qy      265 CCAATGCTTTGATGATGATTTGATGACACAGACTGTGAAGAAATCGTGAGAGGCA 324
Db      119 ProAspAlaPheAspThrGlnGlnLeuLeuHisCysAlaGlnThrLeuLysSerGlyGln 138
Qy      325 ACGGTGAGGTGCGGACCTATGATTTGTGACACACTCAAGTTTACAGAGAGC---ACG 381
Db      139 ProTyrGlnValProIleTyrAspPheLysThrIleGlnArgAspSerPheArg 158
Qy      382 GTGGCTAACCTCGGACAGTGGTCTGTGTGAGGACATTTGTGTCTACAGCCAGAG 441
Db      159 GlnValAlaAsnAlaSerAspValIleIleLeuGlnGlyLeuValPheHisAspSerArg 178
Qy      442 ATCCGGACATGTTCCACTCGGCTCTTCTGTGACACCGACTCCGAGCTGAGCTGTCT 501
Db      179 ValArgAsnLeuMetAsnMetLysIlePheValAspThrAspAlaAspValArgLeuAla 198
Qy      502 CGAAGAGTTTCCCGGAC---GTGCGCGGAGGAGGAGGAGCTGTGAGACAGATCTTACAGC 558
Db      199 ArgArgIleArgAlaGlyAspThrValGlnArgGlyArgAspValAlaAsnSerValLeuGln 218
Qy      559 TACACCACTTGTGAAGCCGCGCTTGTGAGAGATTTGTGCTGCGCAGCAAGAAAGTATGCC 618
Db      219 TyrAlaLysPheValLysProAlaPheAspPheValLeuProSerLysTyrAla 238
Qy      619 GATGTGATCATCCACAGAGAGTGGACAAATGTGTGTCATCAACCTGATGTGTGACAC 678
Db      239 AspValIleIleProArgGlyGlyAspAsnHisValAlaValAspLeuIleThrGlnHis 258
Qy      679 ATCCAGGACATTTG---AATGTGACATCTGCAAAATGCGACGAGAGGGTCCATGGG 735
Db      259 IleHisThrLysLeuGlnGlnHisAspLeuCysLysIleTyr-----ProAsnVal 275
Qy      736 CGAGACTACAGCGGACCTTTTGTGAGCGAGGAGCAAC 774
Db      276 TyrValIleGlnSerThrPheGlnIleArgGlyMetHis 288

RESULT 5
ID      09LTY6      PRELIMINARY;      PRT;      465 AA.
AC      09LTY6;
DT      01-OCT-2000 (TREMblrel. 15, Created)
DT      01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT      01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE      Uridine kinase-like protein.
OS      Arabidopsis thaliana (mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosoids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Columbia;
RA      Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL      Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Columbia;
RX      MEDLINE=20277480; PubMed=10819329;
RA      Nakamura Y.;

```

```

RT      "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT      features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT      clones."
RL      DNA Res. 7:131-135 (2000).
DR      EMBL; AB024028; BA95720.1; -.
DR      HSSP; Q2698; 1BD3.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR000764; Uridine_kin.
DR      Pfam; PF00485; PRK.1
DR      PRINTS; PR00988; URIDINKINASE.
DR      TIGRFAMs; TIGR00235; udk; 1.
KW      Kinase.
SQ      SEQUENCE 465 AA; 52125 MW; 4CD82DB9059E7168 CRC64;

Alignment Scores:
Pred. No.:      8.3e-41      Length:      465
Score:          568.00      Matches:      115
Percent Similarity: 66.94%      Conservative: 51
Best Local Similarity: 46.37%      Mismatches: 68
Query Match:    37.52%      Indels:      14
DB:              10      Gaps:        5

US-09-896-522-3 (1-834) x 09LTY6 (1-465)
Qy      37 GCGCGGAGGCGGACCGCTCGACACAGCGGCTTCTGATAGAGGTGAGCGGCGGACT 96
Db      19 SerProSerAlaProAlaProLeuLysGlnProPheValIleGlyValAlaGlyLysThr 38
Qy      97 GCCAGCGGAGAGTGCACCGTGTGTGAAAGATCATGAGTGTGTGGACAGAAAGAGTGTG 156
Db      39 AlaSerGlyLysThrThrValCysAsnMetIleMetSerGlnLeu----- 53
Qy      157 GAACACGCGGAGGAGAGGTGTATCTCGAGCAGACAGAGTTCACAAAGTCTGACG 216
Db      54 -----HisAspIleArgValAlaLeuValAlaAsnGlnAspSerPheTyrHisSerLeuThr 71
Qy      217 GCAGACGACAGAGCGCAAGGCTTGAAGAAGACAGTAAATTTGACCATCCAGATGCTTT 276
Db      72 LysGlnLysLeuAsnLysVal-----HisGlnTyrAsnPheAspHisProAspAlaPhe 89
Qy      277 GATTAATGATTTGACAGAGACTGTGAAGAAATCATGTGAGAGGCGAAACGGTGAAGTGTG 336
Db      90 AsnThrGlnValLeuLeuSerCysMetGlnLysLeuArgSerGlyIleProValAlaAsnIle 109
Qy      337 CCAGACTATGATTTTGTGACACACTGACAGTTTACAGACCAAGGCTGTACCGTGGG 396
Db      110 ProSerTyrAspPheLysIleHisGlnSerIleGlnSerSerSerProValAsnProGly 129
Qy      397 GACGTGTTCTGTGAGGACATCTTGTGTTCTACAGCCAGAGATCCGCGGACATGTTG 456
Db      130 AspValIleIleLeuGlnGlyLeuValAlaAsnAspProArgValArgAspLeuMet 149
Qy      457 CACCTGGCCTCTTGTGACACCGACTCCGAGCTGAGGCTGTGTGAAGAGTTCTCCGG 516
Db      150 AsnMetLysIlePheValAspThrAspAlaAspValArgLeuSerArgArgIleGlnArg 169
Qy      517 GAC---GTGCGCGGAGGAGGAGCTGAGACAGATTTGACAGAGTTCACCACTCTGG 573
Db      170 AspThrValGlnArgGlyArgAsnIleGlnAsnValLeuGlnIleThrLysPheVal 189
Qy      574 AACCGGCTTGTGAGAGTCTGCTGCGACCAAGAAAGTATGCGCATGATCATCCCA 633
Db      190 LysProSerPheAspLysThrIleGlnProSerMetLysTyrAlaAspIleIlePro 209
Qy      634 CAGAGAGTGACAAATGTGTGTCATCAACTGTATGTGTGACAGACATCCAG---GACATT 690
Db      210 ArgGlyGlyAspAsnAspValAlaIleAspLeuIleValGlnHisThrLysLeu 229
Qy      691 CTAAATGTGACATCTGCAATGTGACACCGAGAGGTTCCAAATGGGCGAGCTCAAGCGG 750
Db      230 CysGlnHisAsnLeuCysLysIleTyr-----SerAsnIlePheIleIleSerSer 246
Qy      751 ACCTTTTGTGAGCGAGGAGGACAC 774

```

Db 247 ThrpneIgmIleuysgllymethis 254

RESULT 6

09LK34 PRELIMINARY; PRT; 483 AA.

ID 09LK34

AC 09LK34

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)

DE Uridine kinase-like protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Eucosmos II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_Taxid=3702;

OC [1]

PC SEQUENCE FROM N.A.

RA SEQUENCE=Columbia.

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RC MEDLINE=20363099; Pubmed=10907853;

RX Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,655 bp covered by ninety P1, TAC and BAC clones.";

RL DNA Res. 7:217-221(2000).

DR EMBL; AF000381; BAB02114.1; -.

DR HSSP; Q26998; IBD3.

DR InterPro; IPR006083; PRK URK.

DR InterPro; IPR00764; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PRO0988; URIDINKINASE.

DR TIGRFAMs; TIGR00235; udk; 1.

Kinase.

SW SEQUENCE 483 AA; 54210 MW; 2862F7AED187020F CRC64;

Alignment Scores:

Pred. No.:	2,296-40	Length:	483
Scores:	563.00	Matches:	117
Percent Similarity:	67.08%	Conservative:	46
Best Local Similarity:	48.15%	Mismatches:	64
Query Match:	37.19%	Indels:	16
DB:	10	Gaps:	7

US-09-896-522-3 (1-834) x 09LK34 (1-483)

0Y CCGCAGCAGCGGCGCCCTTCCGTATGAGGGGTGAGCGCGGCGACATGCGAGCGGAGAGTCGACC 114

|||||

Db 59 ProhneIgmIleuysgllymethis 77

115 GTGTGTGAGAGATCATGTGAGTTGCTGGACAGAAAGAGTGGAAACAGCGGACCGGAG 174

|||||

Db 78 ValCysAspMetIleIleGlnGlnLeu-----HisAspHisArg 90

175 GTGTGATCTCTGAGCGGACGAGGTTCTCAAGGTCCTGACGCGGACGACGACGACGAGCCAG 234

|||||

Db 91 IleValIleuValIleuGlnAspSerPheTyraArgIleuIleuSerGlnIleuGlnHis 110

|||||

0Y GCCCTTGAAGAGAGTACATTTTTCGATCGATCGATGCGCTTGTATGATGATTTGATGAC 294

|||||

Db 111 Val-----GlnIleuTyraIleuAspHisPheAspHisPheAspHisPheAspHisPhe 128

|||||

0Y AGGACTTGAAGACATCTGTGAGGGGCAAAAGGTGAGAGTTCGACCTATGATTTTGTG 354

|||||

Db 129 HisCysValAspIleuIleuSerGlnIleuProTyraGlnIleuProIleuTyraAspPhe 148

|||||

0Y AACACATCAGGATTTACAGAGAC---AGGAGGTCTACCCGCGGAGCGTGTCTGTGTT 411

|||||

Db 149 ThrHisGlnIleuTyraValAspHisPheArgGlnValAspHisPheArgGlnValAspHisPhe 168

|||||

[illegible]

US-09-896-522-3 (1-834) x Q19583 (1-515)

QY	61	CAGCGGCGCCCTTCTGATAGAGGGGAGCGGCGGACCTGCGACGCGGAAAGTCGACCGGTGT	120
QY	62	LYHH:SPRPHheValIleGlyValCysGlyGlySerIleSeriIystrHrtrValAla	81
QY	121	GAGAAATCATGAGGTGTCTGGGACGAAACGAAGTGAACAGCGGACGCGGAAAGTGTCTC	180
QY	82	GIuYrYbIleValGIuHxYrLeuGly-----IleProtrYValThr	94
QY	181	ATCTTGAGCCAGGACAGGTCTTACAAAGTTCCTTGACGCGACGACGAAAGGCCAAAGCCCTTG	240
QY	95	IleIeuSeMeuAAsperPheYrYrYsValIleuthrProGluGluIleYrYsAlaIaHis	114
QY	241	AAAGGACAGTAAATTTTGACCATTCAGATGCGCTTGATATGATATTTTGATGACGAGACT	300
QY	115	GIuSerArGTrYrAsnPhenAspGlyProAsnAlaPheAspPheAspIeuYrYrGlyuVal	134
QY	301	CTGAAGAACATCTGTGGAGGGCAAAACSGTGAAGTCCGACCTATGATTTTGTGACACAC	360
QY	135	LeuYrYsArGLeuAlaGlyGluGlyYsSerValAspValProValYrYrAspPheAsnThrHis	154
QY	361	TCAAGGTTATCCAGAGCCACGCGGAGTCTACCTCGCGGACGAGTGCTGTGTTGAGGCGATC	420
QY	155	SerYrGAsPProAsnSerYrYsMetCetrYrGlyAlaAspAlaIleuPheGluGlyIle	174
QY	421	TTGGGTGTTCTACAGCCACGAGATCCCGGACATGTTCCACCTGCTCTTGTGACACCC	480
QY	175	LeuAlaPheHisAspGlyuYrGlyIleYrYsAsnIleuMetAspMetYrYsValPheValAspThr	194
QY	481	GACTCCGACAGTCAGGCTGTCTCGAAGATGTTCTCCGGAAGTGGCG--CGAGGAGAGGAC	537
QY	195	AspGlyAspLeuAlaGlyLeuAlaTrGAlaValAlaTrGAspValThrAspArGTrGlyArGAsp	214
QY	538	CTGGAGCAGATTCTGACGACGATACACCACTTCCTGTCAGACCGGCTTCGACGAGTGTCTGC	597
QY	215	IleAspGlyIleMetCgluGluInYrYrPheThrPheValIlyAspProAlaPheAspYrYrIle	234
QY	598	CTGCGGACAAAGAGTATGCCGATGTGATTCATCCACGAGAGTGAACATATGTTGCC	657
QY	235	AlaProCysMetAspSerAlaAspIleuValProArGlyGlyIluAsnAspValAla	254
QY	658	ATCAACCTGATGTCGACGACACATC	681
QY	255	IleAspMetIleValGlnAsnVal	262

RESULT 8

ID	Q9U317	PRELIMINARY;	PRT;	555 AA.
Q9U317	Q9U317			
AC	Q9U317:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	F19B6.1b protein.			
OS	F19B6.1 OR F19B6.1B.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;			
OC	Rhabditidae; Peloderae; Caenorhabditis.			
OX	NCBI_Taxid=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Thomas K.;			
RL	Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology."			
RL	Science 282:2012-2016(1998).			
DR	EMBL; Z69635; CA93462.1; -.			
DR	HSSP; Q26998; 1BD3.			
DR	WormPep; F19B6.1b; CE20700.			

```

DR   InterPro; IPR006082; PRK.
DR   InterPro; IPR006083; PRK_URK.
DR   InterPro; IPR00764; Uridine_kin.
DR   Pfam; PF00485; PRK; 1.
DR   PRINTS; PR00478; PHRIKINASE.
DR   PRINTS; PR00388; URIDIKINASE.
DR   TIGRFAMs; TIGR00235; udk; 1.
SQ   SEQUENCE   555 AA;  62673 MW;  D0786FA9B98EBCF98 CAC64

```

Pred. No.:	5,29e-40	Length:	555
Score:	559.00	Matches:	110
Percent Similarity:	74.04%	Conservative:	44
Best Local Similarity:	52.88%	Mismatches:	46
Query Match:	36.92%	Indels:	8
DB:	5	Gaps:	2

US-09-896-522-3 (1-834) x Q9U317 (1-555)	
QY	61 CAGCGGCGCTTCTGTATAGAGGGTGAACGGCGGACATCGCCAGCGGAAGTGCACCGTGTGT 120
Db	102 LYNH1PRPhVal111eclYValCysglYglYSerAlaSerAlaSerGlySerThrValAla 121
QY	121 GAGAAATATATGAGTGTGTGTGGACAGAACAGAGGTGAACAGCGGACCGGAAGCTGTCT 180
Db	122 GlnLYsIleValAlaGlnArgLeuGly-----IleProTyrValThr 134
QY	181 ATCCGTAGCCAGACAGGTTTTCACAGGCTCTGCAGCGGACAGAGCGCAAGCGCCCTTG 240
Db	135 ILIeueSerMetAcAspSerPheTyrIlyValIleuThrProGluGluIleYsAlaAlaHis 154
QY	241 AAGAGCAGTACAAATTTTGACCATTCAGATCCCTTGAATATGATTTGATGCACAGACT 300
Db	155 GlnSerArgTyrAsnPhenSpolYProAsnAlaPheAspPheAspLeuLeuYrGluVal 174
QY	301 CTGAAGAATATGCTGGAGGGCAAAACGGTGAAGGTGCCCATATGATTTTGTGACACAC 360
Db	175 LeuLYsArgLeuArgGluGluLYsSerValAspValProValTYrAspPheAsnThrHis 194
QY	361 TCAAGGTTACACAGACGACCGGTGTTCACCCCTGCGGACGGCGTGTCTGTTGAGGCATC 420
Db	195 SerArgAspProAsnSerIlyMetMetCtyrGlyAlaAspValLeuIlePheGluGlyIle 214
QY	421 TTGGTGTTCACAGCCAGAGAGATCCGGACATGTTCCACCTGCCTCTTCTGTGACACC 480
Db	215 LeuAlaPheHisAspGluArgIleuYsAlaMetSerMetAspMetIlyValPheValAspThr 234
QY	481 GACTCCGACAGCTAGGCTGTCTGGAAGACTTCTCCGGACGTGGC--CGAGGAGGAC 537
Db	235 AspGluYAspLeuArgLeuAlaArgIleValAlaArgAspValThrAspArgGlyAspArg 254
QY	538 CTGGCAGACAAATTCTGACCGAGTACACCACTTGTGAAGCGCGGCTTCGAGAGTCTGC 597
Db	255 ILAspGluYIleMetGluGlnIlyThrPheThrPheValIlyProAlaPheAspIlyrIle 274
QY	598 CTGCCACAAAGAGATATGCGGATGTGATCTCCACAGAGAGTGCACATATAGTTGGC 657
Db	275 AlaProCYMetAcAspSerAlaAspLeuIleValProArgIlyGluIleAsnAspValAla 294
QY	658 ATCAACCTGATCTGCACGACATC 681
Db	295 ILAspMetIleValGlnAsnVal 302

RESULT 9	
Q8YB2	PRELIMINARY; PRT; 466 AA.
AC	Q8YB2;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Putative uracil phosphoribosyltransferase.
EN	ATIG55810.
OS	Arabidopsis thaliana (Mouse-ear cress).

[illegible]

D	b		140	ArgArg---	-ValAsnProSerAspValIleIleLeuGluGlyIleLeuIlePheHisAsp	150
O	y		436	CAGAGATCCGGAGACATTTCACCTCGCCTTTGTGGACACCAGTCCGACTGACG	495	
D	b		159	ProAlaValAArgAspLeuMetAsnMetLysIlePheValAspAlaAspAlaValArg	178	
O	y		496	CTGATCTCGAAGATTCTCCGGGAC---GTGCCCGAGGAGGAGCACTGGACGAATTCTG	552	
D	b		179	LeuAlaAGAgtIleLysArgAspThrValGluLysGlyArgAspIleAlaThrValLeu	198	
O	y		553	ACGAGATCACACCTTGCTGTGAAGCGGCCTTCGAGAGATTGTGCTGCCGACAAAGAAG	612	
D	b		199	AspGlnTyrserylSphneValLysProAlaPheGluAspPheIleuProThrylsYs	218	
O	y		613	TATGCCGATGTGATCATCCACGAGAGGTGACAATAATGTTGCCATCACTGATCGTG	672	
D	b		219	TyrAlaAspIleIleIlePheProArgLyGlyAspAsnHISvalAlaIleLeuIleVal	238	
O	y		673	CAGACATCCAGGACATTCTG---AATGTGACATCTGCAATGGCACCAGAGAGGTCC	729	
D	b		239	GlnHisIleHisThrLysLeuGlyGlnHisAspLeuCyLSylsIleTy-----Pro	255	
O	y		730	AATGGCGGAGACTCAACAGCGGACCTTTTCTGTAGCCAGCGAGAACACCTCGAGTGCAC	789	
D	b		256	AsnLeutyrtValIleGlnInserrThrPheGlnIleArgLyGlyMechisThrLeuIleArgAsp	275	
O	y		790	TCTGGCAACCGGTCACAT 807		
D	b		276	SerLysThrThrLysHis 281		
 RESULT 10 Q9V810 PRELIMINARY; PRT; 614 AA.						
I	D		Q9V810			
A	C		Q9V810; Q9V811;			
D	T		01-MAY-2000 (TREMBLrel. 13. Created)			
D	T		01-MAY-2000 (TREMBLrel. 13. Last sequence update)			
D	T		01-OCT-2002 (TREMBLrel. 22. Last annotation update)			
D	E		CG4798 protein.			
G	N		CG4798.			
O	S		Drosophila melanogaster (Fruit fly).			
O	C		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
O	C		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
O	X		Ephydroidea; Drosophilidae; Drosophila.			
R	N		[1]			
R	P		SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).			
R	C		STRAIN=BERKELEY;			
R	K		MEDLINE=20196006; PubMed=10731132;			
R	A		Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
R	A		Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
R	A		George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
R	A		Sutton G.G., Wortman J.R., Yeandlel M.D., Zhang Q., Chen L.X.,			
R	A		Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
R	A		Abriil K.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
R	A		Ballew R.M., Baasari A., Baxendale J., Bayraktaroglu I., Beasley E.M.,			
R	A		Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
R	A		Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
R	A		Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
R	A		Cherry J.M., Clawley S., Dahlke C., Davenport L.B., Davies P.,			
R	A		De Paolis B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,			
R	A		Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
R	A		Dubin K.U., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
R	A		Fowler C., Gabrieliian A.E., Gang N.S., Gelbart W.M., Glasser K.,			
R	A		Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
R	A		Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
R	A		Hoeftin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwu C.,			
R	A		Jaislin M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
R	A		Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
R	A		Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
R	A		Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
R	A		Merkulov G., Milshina N.V., Nobarty C., Morris J., Moshrefi A.,			
R	A		Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			

RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sinden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195 (2000).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 DR EMBL; AE003803; AAF57871.1; -;
 DR EMBL; AE003803; AAF57873.1; -;
 DR HSSP; Q26998; 1BD3.
 DR FlyBase; FBgn0034213; CG4798.
 DR InterPro; IPR006083; PRK_UK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRfams; TIGR00235; udk; 1.
 KM Alternative splicing; Hypothetical protein.
 FT VARSPLIC 1 207 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 614 AA; 68704 MW; C81B6B3B9D392A1 CRC64;

Alignment Scores:
 Pred. No.: 1.65e-38 Length: 614
 Score: 542.00 Matches: 106
 Percent Similarity: 71.84% Conservative: 42
 Best Local Similarity: 51.46% Mismatches: 50
 Query Match: 35.80% Indels: 8
 DB: Gaps: 3

US-09-896-522-3 (1-834) x Q9V810 (1-614)

QY CCCTTCCTGATAGGGGTAGCGGCGGACCTGACGGGAAAGTGCACCGTGTGAGAG 126
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 174 ProheValIleGlyIleCySgIySerIaSerIySlnrThrValAlaGlnuys 193
 QY 127 ATCATGAGTGTCTGGACAGAACAGAGTGGACAGCGGACCGAAAGTGTGATCTG 186
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 194 IleIleGlnSerIeu-----AspValProTrp-----ValThrLeuLeu 206
 QY 187 AGCCAGACAGAGTCTTCAAGAGTCTTCAAGAGAGCAAGCAAGCAAGCTTGAAGA 246
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 207 SerMetAspCysPheTyrluysIleLeuAsnGlnuysGlnIleGlnAlaLeuIleAsn 226
 QY 247 CAGTACAAATTTTGAACATCCAGATGCTTGTATGATGTTGATGACACAGGACTGTGAG 306
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 227 GlnuysrAsnPhenAspIleProAspIlePhenAspIleGlnLeuLeuAspValLeuThr 246
 QY 307 AACATGTGAGGGGCAAAACGCTGAGAGTGGCGGACCTATGATTTTGTGACACTCAAG 366
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 247 LysLeuysGlnuysIyAryGlyValGlnuysAlaProvalTyrlAsnPhenValThrIleAsnGlyAryG 266
 QY 367 TTACCAAGACACAGCGGTGTCTACCTCGGACAGTGTCTGTTGAGGCGACTTTGGTG 426
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 267 GluSerGlnThrIySlnrMetTyrlAlaAsnValIleIlePheGlnuysIleLeuThr 286
 QY 427 TTTCACAGCCAGAGATCCGGACATGTCATCCACTGCGGCTTCTGTCGACACCGACTCC 486
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 287 PheIleSerProGlnuysValLeuLeuysLeuAspMetCysIleIlePheValAspThrAspPro 306
 QY 487 GAGCTCAGCGTGTCTGACAGATTCTCCGGACGCTG--CGCCAGAGAGAGCACTTGAG 543
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 307 AspIleAryGlnuysAlaAryGlnuysAryGlnuysIleSerGlnuysIyAryGlnuys 326
 QY 544 CAGATTCTGACGACGATACCACTTGTGAGAGCGGCTTTCAGAGAGCTTCTGCTCCG 603
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 327 GlnuysValLeuysGlnuysIyLeuAsnMetValIySlnrProSerTyrlCyAsnTyrlIleAlaPro 346

QY 604 ACAAGAAGTATCCGATGATCATCCACGAGAGTGCAGATATGTTGCATCAAC 663
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 347 ThrMetAlaIleAspIleIleValProAryGlnuysAlaAspAsnIyValAlaIleIleHis 366
 QY 664 CTGATCGTGCAGACATC 681
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 367 LeuIleValGlnuysIleVal 372

RESULT 11

Q8WQK4 PRELIMINARY; PRT; 626 AA.
 AC Q8WQK4;
 DT 01-OCT-2002 (TEMBLrel. 22, Created)
 DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE L003595P.
 GN CG4798.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Friese E.,
 RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuroo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY129436; AAM76178.1; -;
 DR FlyBase; FBgn0034213; CG4798.
 DR InterPro; IPR006083; PRK_UK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRfams; TIGR00235; udk; 1.
 SQ SEQUENCE 626 AA; 70001 MW; 9DBEC0A7331A6EF3 CRC64;

Alignment Scores:

Pred. No.: 1.65e-38 Length: 626
 Score: 542.00 Matches: 106
 Percent Similarity: 71.84% Conservative: 42
 Best Local Similarity: 51.46% Mismatches: 50
 Query Match: 35.80% Indels: 8
 DB: Gaps: 3

US-09-896-522-3 (1-834) x Q8WQK4 (1-626)

QY 67 CCCTTCCTGATAGGGGTAGCGGCGGACCTGACGGGAAAGTGCACCGTGTGAGAG 126
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 186 ProPheValIleGlyIleCySgIySerIaSerIySlnrThrValAlaGlnuys 205
 QY 127 ATCATGAGTGTCTGGACAGAACAGAGTGGACAGCGGACCGAAAGTGTGATCTG 186
 |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 206 IleIleGlnSerIeu-----AspValProTrp-----ValThrLeuLeu 218
 QY 187 AGCCAGACAGAGTCTTCAAGAGTCTTCAAGAGAGCAAGCAAGCAAGCTTGAAGA 246
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 219 SerMetAspCysPheTyrluysIleLeuAsnGlnuysGlnIleGlnAlaLeuIleAsn 238
 QY 247 CAGTACAAATTTTGAACATCCAGATGCTTGTATGATGTTGATGACACAGACTTGAG 306
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 239 GlnuysrAsnPhenAspIleProAspIlePhenAspIleGlnLeuLeuAspValLeuThr 258
 QY 307 AACATGTGAGGGGCAAAACGCTGAGAGTGGCGGACCTATGATTTTGTGACACTCAAG 366
 ::|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
 Db 259 LysLeuysGlnuysIyAryGlyValGlnuysAlaProvalTyrlAsnPhenValThrIleAsnGlyAryG 278
 QY 367 TTACCAAGACACAGCGGTGTCTACCTCGGACAGTGTCTGTTGAGGCGACTTTGGTG 426


```

Db      120 LeuArgysclglnAlaValaAspIleProAsnIlyrAspHeuysSserryLyAsaAsn 139
Qy      367 ---TTACAGAGACCAACGCGTGTACCCGCGAGCTGCTGTGTTGAGGCGATCTTG 423
Db      140 ValPheProValArgArg---ValaAsnProSerAspValIleIleLeuGlnGlyIleLeu 158
Qy      424 GTGTTCTACAGCCAGAGAGATCCCGAGATGTTCCACCTGCGCCCTTCTCGTGAACACC--- 480
Db      159 IlePheHisAspProValArgAspLeuMetAsnMetLysIlePheValaAspAlaGly 178
Qy      480 ----- 480
Db      179 LeuSerHisThrIlyrProValaAsnThrTyrValValysSerValaIaIyMetArgArg 198
Qy      481 -----GACTCCGAGCTGAGCTGTCTGCAAGAGTTCTCCGG 516
Db      199 CysThrCysIleCysThrHisGlnAspAlaAspValaIArgLeuAlaArgIleLysArg 218
Qy      517 GAC---GTGGCGGAGGAGGAGGAGCTGGAGAGATTTGTAGACGATACACCACTTGTG 573
Db      219 AspThrValGlnLysGlnArgAspIleAlaThrValaLeuAspGlnIlyrSerLysPheVal 238
Qy      574 AAGCGCGCTTCCAGAGATTCTGCGCTCCGACAAGAAGATGCGGATGATCATCCCA 633
Db      239 LysPheAlaPheGlnAspPheIleLeuProThrIlyrLysTyrAlaAspIleIlePro 258
Qy      634 CGAGAGATGACATATGTTGCTGATCACTGATCTGTCAGACATCCAGACATTTCTG 693
Db      259 ArgGlyGlnAspAsnHisValaIaIleAspLeuIleValaGlnHisIleHisThrLysLeu 278
Qy      694 ---AATGTCGACATCTGCAATATGSCACCGAGAGGCTCAATGGCGGAGCTCAACCGG 750
Db      279 GlnGlnHisAspLeuCysLysIleTyr-----ProAsnLeuTyrValaIleGlnSer 295
Qy      751 ACCTTTCTGAGCGAGCGAGCACTGAGATGCTGATCTGTCGCAACGCTGACAT 807
Db      296 ThrPheGlnIleArgGlyMetHisThrLeuIleArgAspSerLysThrThrLysHis 314

RESULT 15
Q9LEF22 PRELIMINARY;  PRF; 1060 AA.
AC Q9LEF22;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE F20N2.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_taxid=3702;
RN
RP SEQUENCE FROM N.A.
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA Shim P., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vayberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
RT 1.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
[4]

```

```

RP SEQUENCE FROM N.A.
RA Shim P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,
RA Walker M.M., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
RA Gonzalez A.A., Hansen N.N.P., Huizar L.L., Kremenevskata I.I.,
RA Lenz C.C., Li J.J., Liu S.S., Luo S.S., Rowley D.D., Schwartz J.J.,
RA Toriumi M.M., Vayotskaya V.V., Yu G.G., Davis R.R.W.,
RA Federspiel N.N.A., Theologis A.A., Ecker J.R.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Cheuk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chio J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vayberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002328; AAF79498.1; -.
DR HSSP; Q26998; 1BD3.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 1060 AA; 116160 MW; 0398B256AB31D6B CRC64;

Alignment Scores:
Score: 2.61e-35 Length: 1060
Percent Similarity: 506.00 Matches: 109
Best Local Similarity: 59.18% Conservative: 49
Query Match: 40.82% Mismatches: 59
Indels: 33.42% Gaps: 50
Gaps: 9

US-09-896-522-3 (1-834) x Q9LEF22 (1-1060)
Qy      31 AGCCCGCGCGCGAG-----GCCGACCGTCCGACAGCGCCCTTCTGATA--- 78
Db      29 SerAsnArgProGlnMetAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 48
Qy      79 -----GGGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 129
Db      49 GluSerArgGlyValaIaGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 68
Qy      130 ATGAGTGTCTGAGGAGACAGACGATGGAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 189
Db      69 MetGlnGlnLeu-----HisAspGlnArgAlaValaValaValaValaValaValaValaVala 81
Qy      190 CAGAGAGGTTCTTACAGGCTCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 249
Db      82 GlnAspSerPheThrHisAsnValaAsnGlnValaGlnLeuValaArgVala-----HisAsp 99
Qy      250 TACATTTTACCATCATCATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309
Db      100 TYAsnPheAspHisProAspAlaPheAspThrGlnGlnLeuLeuSerSerMetGluLys 119
Qy      310 ATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366
Db      120 LeuArgLysGlnAlaValaAspIleProAsnIlyrAspPheLysSerTyrLysAsaAsn 139
Qy      367 ---TTACAGAGACCAACGCGTGTACCCGCGAGCTGCTGTGTTGAGGCGATCTTG 423
Db      140 ValPheProValArgArg---ValaAsnProSerAspValIleIleLeuGlnGlyIleLeu 158
Qy      424 GTGTTCTACAGCCAGAGAGATCCCGAGATGTTCCACCTGCGCCCTTCTCGTGAACACC--- 480
Db      159 IlePheHisAspProValArgAspLeuMetAsnMetLysIlePheValaAspAlaGly 178
Qy      480 ----- 480

```

```
Db      179 LeuSerHisThrLysProValAsnThrTyrValValLysSerValAlaTyrMetArgArg 198
QY      481 -----GACTCCGACGTCAGGCTGTCTCCGAAGATTCTCCGG 516
Db      199 CysThrCysIleCysThrHisGluAspAlaAspValArgLeuAlaArgArgIleLysArg 218
QY      517 GAC--GTGCGCCGAGGAGGAGGACCTGAGCAGATTCTGACGACATACACCACCTTCGTG 573
Db      219 AspThrValGluLysGlyArgAspIleAlaThrValLeuAspGlnTyrSerLysPheVal 238
QY      574 AAGCCGCTTCGAGGAGATTCTGCTGCCGACAAAGATATGCCGATGTGATCATCCCA 633
Db      239 LysProAlaPheGluAspPheIleLeuProThrLysTyrAlaAspIleIleIlePro 258
QY      634 CGAGGAGTGCACATATGTTGCCATCAACCTGATCGTCAGACATCCAGACATTCTG 693
Db      259 ArgGlyGlyAspAsnHisValAlaIleAspLeuIleValGlnHisIleHisThrLysLeu 278
QY      694 --AATGGTGACATCTGCATA 711
Db      279 GlyGlnHisAspLeuCysLys 285
```

Search completed: November 25, 2003, 07:39:11
Job time : 69.1806 secs

RESULT 1	UCLK1_HUMAN	STANDARD;	PRT;	277 AA.
AC	09HA47;			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCLK 1) (Uridine monophosphokinase 1) (Cytidine monophosphokinase 1).			
GN	UCLK1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A. AND CHARACTERIZATION.			
EX	MEDLINE=21203813; PubMed=11306702;			
RA	Van Rompay A.R., Norda A., Lindén K., Johansson M., Karlsson A.;			
RT	"Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases.";			
RL	Mol. Pharmacol. 59:1181-1186(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Ho Y.S., Johnson R.K.;			
RT	"Human uridine kinase from prostate cancer cell line (LMCap).";			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUES=Mammary gland;			
RA	Iisogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Masatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Matsumoto S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nagahata K., Masuno Y.,			
RA	Niimiya K., Iwayanagi T.;			
RT	"NDO human cDNA sequencing project.";			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Xin Y.R., Yu L., Zhao S.Y.;			
RT	"Cloning of a new human cDNA similar to Mus musculus uridine kinase			

```

RT mRNA ";
RU Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-antioleylcytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF237290; AAK8334.1; -.
DR EMBL: AF254133; AAK49122.1; -.
DR EMBL: AK022317; BAB14010.1; -.
DR EMBL: AF125106; AAL75943.1; -.
DR InterPro: IPR006083; PRK_UK.
DR InterPro: IPR007644; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS: PR00988; URIDINKINASE.
DR TIGRfams: TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DCEPAPAD -> GARARAGAN (IN REF. 4).
FT CONFLICT 56 57 OR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MM; AFD9ED92780CD502 CXC64;

Alignment Scores:
Pred. No.: 1,1e-101 Length: 277
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.77% Indels: 0
Gaps: 0

US-09-896-522-3 (1-834) x UCK1_HUMAN (1-277)
QY 1 ATGGCTTCGGGGGAGGCGAAGATCGGAGCGCCCGCGGAGCGCGACCGTCCGCAC 60
DB 1 Metlaserlaaglygluaspysgluserproalaprogluilaasparpghis 20
QY 61 CAGGGCGCCCTTCGATAGGGGGTGAAGCGGCGACTCGACGCGGAAGTGCACGCTGTG 120
DB 21 GlnhrgproheliuegllyalsercllylthralaserglyluserthvalCys 40
QY 121 GAGAAATCATGAGATTGCTGGACAGAACGAGGTGGAACGCGGACGCGAAGTGTG 180
DB 41 GlulysilemetgluleuenglglahsmgluvalgluhlnaglnaglyvalVal 60
QY 181 ATCTGAGCCGAGGACAGTTCTACAAGTCTCGACCGCAGAGCGAAGGCCAGGCTTG 240
DB 61 lIleuserglinaapargpnehyrlyvalleuthralaglnuglnlysalalysaleu 80
QY 241 AAAGAGATGATATTTGACATCCAGATCCCTTTGATATATGATTGATGACAGGACT 300
DB 81 llyselylntlyraanpneasphehisprohpalalaphesapnshpdeumethisargthr 100
QY 301 CTGAAGAACATCGTGAGAGGCGAAACCGGTGAGGTGCCGACTTATGATTTTGTGACAC 360
DB 101 leuysaenilevalgluglyllyethrValgluValProthrlyraapnevalthhis 120

```

```

QY 361 TCAAGGTTACCAAGACCAACGCGTGTACCCCTCGCGGACGTGGTCTGTTTGAAGGCGATC 420
DB 121 SerAyleupProdlunhrthrValValIyrProalaAspValValIeuphegluglylle 140
QY 421 TTGGTTTTCACAGCCAGAGATCCGAGACATGTTCCACTCGGCTCTTCTGACAC 480
DB 141 LevalPheylserlntglunlleargapmetPhehisleuayrgueuhelalaprthr 160
QY 481 GACTCCGAGTCGAGCGTGTCTCGAAGATTTCTCCGGACGTGGCGCCGAGGAGACCTG 540
DB 161 AspserAspValalarguserhargValIleuayrgAspValalargylargAspLeu 180
QY 541 GACCAATTCGACGAGATCAACCACTTCGTAACCGCGGCTTCGAGAGATTCTGCTCG 600
DB 181 GlnglnlleuthrlntrlntrPhevalLysproalalphegluglnpneCysleu 200
QY 601 CCGACAAAGATATCCGATGTGATCATCCACAGAGAGTGCACATATGTGTCATC 660
DB 201 ProthrlybelysryrlyalaspvalilleleprohrglyValalaspnmetvalalaile 220
QY 661 AACCTGATGTCGACGACATCCAGACATTTCTGAATGTGACATCTGCATAATGCCAG 720
DB 221 AsnleuilevalglinhisrleglnaepilleuasnlyAspilleCyslystrpHisarg 240
QY 721 GAGAGGTTCGAATGGGCGAGCTACAGAGGACCTTTTCGAGCGAGGAGACACCTTG 780
DB 241 GlyglySerasnlyllyrserlyrlylsargthrPheSerGlnProdllyaspHisprogly 260
QY 781 ATGCTGACCTCTGGCAACGCTGACATTTGAGTCCAGAGAGACCCGAC 831
DB 261 MetleuthrsercllylsarhsrhisleuagluserSerArygrhHis 277

RESULT 2
UCK1_MOUSE STANDARD; PRT; 277 AA.
ID UCK1_MOUSE
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Caesvant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Wulshy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fehey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.

```



```

CC -1- TISSUE SPECIFICITY: Expressed in placenta.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF236637; AAK14053.1; -.
DR EMBL; AB062451; BAB56162.1; -.
DR Genew; HGNC:12562; UMPK.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR007644; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding.
DR NP BIND 27
SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;

Alignment Scores:
Pred. No.: 7.67e-63 Length: 261
Score: 934.00 Matches: 189
Percent Similarity: 79.41% Conservative: 27
Best Local Similarity: 69.49% Mismatches: 40
Query Match: 61.69% Indels: 16
DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x UCK2_HUMAN (1-261)
QY 22 GACTGCAGAGCCCCGCGCGCGAGGCGACCCGCGAC---CAGCGGCGCTTCTGATA 78
DB 4 AspergillusnigerGlnAAsnHnGlnInProAsnGlyGlnProPheLeuIle 23
QY 79 GGGGTGAGCGGCGGCGACTGCGAGCGGGAAGTGCACCGTGTGTGAAAGATCATGAGTTG 138
DB 24 GlyAlaSerGlyGlnThrAlaSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43
QY 139 CTGGGACAGAACGAGGTGGAACAGCGGCGGAAAGGTGTCATCCGTGAGCAGACAGG 198
DB 44 LeuAlaGlnAsnGlnValAspTyrArgGlnLysGlnValAlaIleLeuSerGlnAspSer 63
QY 199 TTCTAAGAGTCTGACGCGAGCAGACAGGCGCAAGGCGCTTGAAGAAGACAGTACATTTT 258
DB 64 PheTyrArgValLeuThrSerGlnGlnValAlaLysAlaLeuLysGlnPheAsnPhe 83
QY 259 GACCATCGAAGTCTTGATTAATGATTGATGACAGAGACTGGAAGACATGCTGAGG 318
DB 84 AspisProAspAlaPheAspAsnGlnLeuLeuLysThrLeuLysGlnIleThrGln 103
QY 319 GCGAAACGCGTGAAGTCCGACCTATGATTTTGTGACACACTGAAGTTACGAGAAC 378
DB 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGlnIleThr 123
QY 379 ACGGTGCTACCTGCGGAGCGTGGTCTGTTGAGGAGCATCTTGTTCTTACAGCCAG 438
DB 124 ValThrValTyrProAlaAspValAlaLeuPheGlnGlnIleLeuAlaPheTyrSerGln 143
QY 439 GAGATCCGGGACATGTTCCACTCGCGCTTCTGTGACACCGACCTCCGAGCTGAGCTG 498
DB 144 GluValAlaArgAspLeuPheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeu 163
QY 499 TCTTGAAGACTTCTCCGGGAGCTG---CGCGGAGGAGGAGCTGAGAGCAGATTCTGACG 555
DB 164 SerTyrArgValLeuAlaArgAspIleSerGlnTyrGlnTyrArgAspLeuGlnIleLeuSer 183
QY 556 CAGTACACACACCTTCTGTGAACCGCGCTTCAAGAGATTCTGCGCGACAAAGATAT 615

```

```

DB 184 GlnTyrIleThrPheValLysProAlaPheGlnGlnPheCysLeuProThrLysLysTyr 203
QY 616 GCCGATGTGATTCATCCACGAGAGTGCACATATGCTTCCATCACTGATCGCGAG 675
DB 204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln 223
QY 676 CACATCAGAGACATTTGTGATGTGACATCTGCAATGACCGAGAGGTCGCAATGGG 735
DB 224 HisIleGlnHisPheIleLeuAsnGly-----Gly 232
QY 736 CGAGCTAACAGCGGACCTTTTCTGAGCCAGGAGACCACTGGAGTGCATCTGGC 795
DB 233 ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrThrProSerArg 249
QY 796 AAACGCTCATTTTGGATTCACGACACAGACCCAC 831
DB 250 LysArgGlnAlaSerGlnSerSerArgProHis 261

RESULT 4
UCK2_MOUSE STANDARD; PRT; 261 AA.
ID UCK2_MOUSE
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
DE monophosphokinase 2) (Cytidine monophosphokinase 2).
GN UCK2 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompey A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RL Mol. Pharmacol. 59:1181-1186(2001).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF236636; AAK14052.1; -.
DR MGD; MGI:1931744; Uck2.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR007644; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding.
DR NP BIND 27
SQ SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;

```

Alignment Scores:

Pred. No.:	1,09e-62	Length:	261
Score:	932.00	Matches:	186
Percent Similarity:	78.31%	Conservative:	27
Best Local Similarity:	68.38%	Mismatches:	43
Query Match:	61.56%	Indels:	16

DB:	1	Gaps:	4
-----	---	-------	---

US-09-896-522-3 (1-834) x UCK2_MOUSE (1-261)

```

QY 22 GACTGCGAGAGCCCGCGCGAGAGCCGACCGTCCGAC---CAGCGGCCCTTCTGTATA 78
   |||
Db 4 AspSerGluGlnThrLeuGlnAmHieGlnGlnProAsnGlyGlyGluProPheLeuIle 23
QY 79 GGGGTGAGCGCGGCATCTGCCAGGGAAGTCCAGCCTGTGTGAGGAATCATGAGAGTTG 138
   |||
Db 24 GlyValSerIlyGlyThrAlaSerGlyLysSerSerValYsAlaIlyValIleGlnLeu 43
QY 139 CTGGGACAGAGCGAGTGGAGAGCGGAGGAGGAGTGTGATCTTGAGCCAGGACAGG 198
   |||
Db 44 LeuGlyGlnAsnGlnValIlePyrThrIleGlnIleValIleLeuSerGlnAspSer 63
QY 199 TTCTACAGAGTCTCTGAGCGAGAGAGAGGCGCCTTGAAAGGACAGTACAATTTT 258
   |||
Db 64 PheTyArgValLeuThrSerGluGlnIlyAlaIlyAlaLeuIlySglIlyGlnPheAsnPhe 83
QY 259 GACCATTCAGATGCTTGTATGATTTGATGACAGAGACTGTGAAGAACATGTGTGAG 318
   |||
Db 84 AspHisProAspAlaPheAspAsnGlnLeuIlePheIlyThrLeuIlySglIleThrGlu 103
QY 319 GGGAAAAGCGTGGAGCGGACCTATGATTTGTGACACAGCTCAAGATTACAGAGAGCC 378
   |||
Db 104 GlyIlyThrValGlnIleProValIlyAspPheValaSerHisSerArgIlySglIlyThr 123
QY 379 ACGGTGCTACCTCTGCGAGAGCGTGTCTGTGTGAGGCACTTGTGTCTACAGCCAG 438
   |||
Db 124 ValThrIleThrProAlaAspValIleValIlePheGluIlyLeuAlaPheTyIlySerGln 143
QY 439 GAATTCGGGAGACATGTTCCACCTGCGCTCTTGTGACACCGACTCCGACGTGAGCTG 498
   |||
Db 144 GluValAlaArgAspLeuPheGlnMetIlyLeuPheValaAspThrAspAlaAspThrArgLeu 163
QY 499 TCTCGAAGAGTCTCTCGGAGAGCGT---CGCGAGAGGAGGAGACCTCGAGAGATTCTACG 555
   |||
Db 164 SerArgThrValLeuAspAspIleSerGlnArgGlyArgAspLeuGlnIleLeuSer 183
QY 556 CAGTACACCACTTCTGTAAGCGCGCTTGTGAGAGATTCTGCTGCGACCAAGAAAGTAT 615
   |||
Db 184 GlnTyIleThrPheValIlySerProAlaPheGluGlnIlyPheCysLeuProThrIlySlyTyr 203
QY 616 GCGGATGTGATCCCGAGAGAGTGTGATGTTGTCATCAACCTGATGTGCGAG 675
   |||
Db 204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValaIleAsnLeuIleValGln 223
QY 676 CACATTCAGAGCACTTGTGAATGTGATGTCGCAATGCGAAGGAGAGGAGGTCGAATGG 725
   |||
Db 224 HisIleGlnAspIleLeuAsnGlyGlyLeuSerIlyArgGlnThrAsnGly----- 240
QY 736 CGGAGCTACAGGAGGAGCCTTTTCTGAGCCAGGGGACCACTGGGATGTGACTCTGGC 795
   |||
Db 241 -----TyrLeuAsn-----GlyTyrThrProSerArg 249
QY 796 AAACGGTACATTGGAGTCCAGAGAGACCCAC 831
   |||
Db 250 LysArgGlnAlaSerGlnIlySerSerArgProHis 261

```

RESULT 5

UCK_DROME	STANDARD;	PRT;	260 AA.
ID_UCK_DROME			
Q9VCG9;			
DT 28-FEB-2003 (Rel. 41, Created)			
DT 28-FEB-2003 (Rel. 41, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			

DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwee R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsch K., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleby J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Styriska S., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J., Wu Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -|- PATHWAY: Pyrimidine salvage pathway.
CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL/ AEO03747; AAF56274.1; -
CC FLYBase; FBgn0039179; CG6364.
CC InterPro; IPR006083; PRK URK.
CC InterPro; IPR00764; uridine_kin.
CC Pfam; PF00485; PRK_1.
CC PRINTS; PR00986; URIDINKINASE.
CC Hypothetical protein; Transferase; Kinase; ATP-binding.
KW NE BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2BE57423704925E6 CRC64;

Alignment Scores:

Pred. No.: 4,38e-48 Length: 260
 Score: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 48.75% Indels: 2
 DB: 1 Gaps: 2

US-09-896-522-3 (1-834) x UCK_DROME (1-260)

QY 67 CCCTCTGATAGGGGTGAGCGGCGGCACTCCGACCGGGAAGTCGACCGCTGTGTGAAG 126
 Db 27 ProthelutleuilegylvalalaglyglYthrIaserglysserThrValCysleuLys 46
 QY 127 ATCATGAGTTCGTCGGGACAGAGAGGTGAACAACGCGGACCGGAAGTGTGATCCTG 186
 Db 47 IlemetGuglnleuilegylvalalaglnuetsphsthrGlnaGlnValaValSerIle 66
 QY 187 AGCCAGGACAGGTTCTACAGAGTCTGACGCGACAGAGGACGACGACGCTTGAAAGA 246
 Db 67 SerGlnaspserspethrYarGlnleuthrProalagluYsalalYsalaglnleu 86
 QY 247 CAGTACAAATTTGACATCCAGATGCTTGTATATATGATTGTGACAGAGACTCTGAAG 306
 Db 87 LeuhenpneasphtsProaspAlaPheasnGlnleuMetYrserThrleuGln 106
 QY 307 AACATCGTGGAGGCAAAACGATGAGGAGGCGGACCTGATTTGTGACACACTCA--- 363
 Db 107 AsnIleuYsGlyhslYsValGlnIleProserYrAspYrArGthrAsnSerleu 126
 QY 364 AGGTACGACAGACACGCGTGTACCTTCGCGACGCTGCTTGTGAGGAGCATCTTG 423
 Db 127 AsphegluValenValleuValIleYrProalaspValleuPhegluGlyIleleu 146
 QY 424 GTGTTCTACACCGGACGAGATCCGGACATGTTCCAGCTGCGCTTGTGTGACACCGAC 483
 Db 147 ValPheYrPhePolYsIleArGlnleuPhehslMetYsleuPheValaspThrAsp 166
 QY 484 TCGACGTCGAGGTCGTGCGAAGGTTCTCGGAGCGTG---CGCCGAGGAGGAGCATCTG 540
 Db 167 SeraspThrArgleuAlaArgArGlnValProArgaspIleasnGlnuArgYalArgAspLeu 186
 QY 541 GAGCAGATTTCTGACGACGATACACCACTTCGTGAGCGGCGCTTCGAGAGTTGCTGCTG 600
 Db 187 AspAlaValleuThrGlnTyrMetThrPheValYsProAlaPhegluGlnuPheCysSer 206
 QY 601 CCGCAAGAAGATATGCCGATGTGATCATCCACGAGAGAGTGCACATATGTTGCCATC 660
 Db 207 ProThrYsYsPheAlaAspValIleIleProArgGlyAlaAspAsnThrValaIle 226
 QY 661 AACCTGATCGTCGACGACATCCAGACATCTGCTG 693
 Db 227 AspIleuValIshslleGlyGlnleu 237

RESULT 6

URL1_HUMAN STANDARD; PRT; 548 AA.
 AC O9NWZ5: O9H322;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 GN Uridine kinase-like 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matanabe K., Kumagai A., Itakura S., Yamazaki M., Tachiro H., Ota T.,
 RA Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Hammond D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharvelaish M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.T.,
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Rose M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Skarn R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: AK000524; BAA91230.1; -;
 DR EMBL: AL118506; CAC15497.1; -;
 DR HSSP: Q26998; 1BD3.
 DR Genew: HGNC:15938; URKL1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK-URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRBLKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRfams: TIGR00235; udk.1.
 KW Transferase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD03697E02FE7DB CRG64;

Alignment Scores:

Pred. No.: 1.62e-33 Length: 548
 Score: 545.50 Matches: 111
 Percent Similarity: 64.20% Conservative: 45
 Best Local Similarity: 45.68% Mismatches: 60
 Query Match: 36.03% Indels: 27
 DB: 1 Gaps: 5

US-09-896-522-3 (1-834) x URL1_HUMAN (1-548)

QY 25 TCGGAGAGCCCGCGCGG-----GAGCGGACCGGT 54
 Db 65 CysLysSerGlnuProleuLeuArgThrSerLysArgThrIleYrThrAlaGlyArg 84
 QY 55 CCG-----CACGACGCGCCCTTCTGATAGGAGGTGAGC 87

```

Db      85 ProProtrpTryrAnsglnHISGLYThrGlnSerlySGlnAlaPheAlaIleGlyLeuGly 104
Qy      88 GGGGGGACTGCCAGCGGGGAAGTCAACCGTGTGTGAAGAATCATGAGTTGCTGGAGAC 147
Db      105 G|G|G|S|e|r|A|s|e|r|G|l|y|s|t|h|r|V|a|l|A|a|r|g|e|t|l|e|g|l|a|l|e|u|----- 122
Qy      148 AACGAGGTGAACAGCGCGACCGGAGGTGTGATCTCTGAGCCAGACAGATTGTACAG 207
Db      123 ---AspValProTrp-----ValValLeuLeuSerMetAspSerPheTrpLys 137
Qy      208 GTCTCTGACGACAGACAGAAAGCCAGCCTTGAAGACATCAATTTGACCATCCA 267
Db      138 ValLeuThrGlnGlnGlnGlnGlnAlaAlaHisAsnAsnPheAsnPro 157
Qy      268 GATGCTTTGATATGATTTGATGACACAGACTTGAAGACATCGTGAAGGCGAAACG 327
Db      158 AspAlaPheAspPheAspLeuIleIleSerThrLeuLysLysGlnGlyLysSer 177
Qy      328 GTGAGAGTGCGGACCTATGATTTGTGACACACTCAAGCTTACACAGACACAGCTGTC 387
Db      178 ValLysValProIleTryrAspPheThrHisSerArgLysAspTrpLysThrLeu 197
Qy      388 TACCTCTCGGACGCTGTCTGTGTGAGGCGCATCTTGCTTCAACGACGAGATCCGG 447
Db      198 TyrGlyAlaAsnValIleIlePheGlnGlyLeMetAlaPheAlaAspLysThrLeu 217
Qy      448 GACATGTTCCACCTGCGCCTTCTGTGTGACACAGCATCCGACGTCGTCTCGAAGA 507
Db      218 GluLeuLeuAspMetLysIlePheValAlaPheAspSerAspIleArgLeuValArg 237
Qy      508 GTTCTCGGAGACGCG---CGCGGAGGAGGACCTGTGACGATCTGACGCGCTACAC 564
Db      238 LeuArgArgAspIleSerGlnArgGlyArgAspIleGlnGlyValIleLysGlnTryrAsn 257
Qy      565 ACCTTCGTGAAGCGCGCTTGAAGAGTTCGCTGCGACGCAAGAAAGTATCCGATGTC 624
Db      258 LysPheValLysProSerPheAspGlnTryrIleGlnProThrMetArgLeuAlaAspIle 277
Qy      625 ATCATCCGACGAGAGTGACATATATGTTGCCATCACTGATCGTGCAGACATCCAG 684
Db      278 ValValProArgLysArgLysThrValAlaIleAsnLeuIleValGlnHisValHis 297
Qy      685 GACATTCGTG 693
Db      298 SerGlnLeu 300

RESULT 7
URL_MOUSE
ID      URL_MOUSE      STANDARD;      PRT;      548 AA.
AC      Q91YL3;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DE      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Uridine kinase-like 1.
GN      URKL1.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22388257; PubMed=12477932;
RA      Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA      Diatchenko L., Mansina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.J., Scheetz T.E.,
RA      Brownstein M.J., Urdin T.B., Toshimuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huljk S.W.,

```

```

RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Whiting J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA      Blakeley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA      Schnerker A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: BC016535; AAH16535.1; -.
DR      WGD; MG1:1915806; 11100710Rik.
DR      InterPro; IPR006082; PRK.
DR      InterPro; IPR006083; PRK URK.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam: PF00485; PRK. 1
DR      PRINTS; PR00476; PRRIDKINASE.
DR      PRINTS; PR00988; URIDINKINASE.
DR      TIGRFAMs; TIGR00235; udk. 1.
KW      Transferase; Kinase; ATP-binding.
FT      NP BIND 105 112 ATP (POTENTIAL).
SQ      SEQUENCE 548 AA; 60841 MW; 99A5CA23B9BD8525 CRC64;

Alignment Scores:
Pred. No.: 4,59e-33 Length: 548
Score: 539.50 Matches: 110
Percent Similarity: 64.20% Conservative: 46
Best Local Similarity: 45.27% Mismatches: 60
Query Match: 35.63% Indels: 27
DB: 1 Gaps: 5

US-09-896-522-3 (1-834) x URL_MOUSE (1-548)
Qy      25 TGCGAGAGCCCGCGCG-----GAGCGGACCGT 54
Db      65 CysLysSerIleuProLeuLeuArgThrSerLysArgThrIleTryrAlaGlyArg 84
Qy      55 CCG-----CACACGCGCCTTCTGTATGAGGCTGAGC 87
Db      85 ProProtrpTryrAnsglnHISGLYThrGlnSerlySGlnAlaPheAlaIleGlyLeuGly 104
Qy      88 GGGGGGACTGCCAGCGGGGAAGTCAACCGTGTGTGAAGAATCATGAGTTGCTGGAGAC 147
Db      105 G|G|G|S|e|r|A|s|e|r|G|l|y|s|t|h|r|V|a|l|A|a|r|g|e|t|l|e|g|l|a|l|e|u|----- 122
Qy      148 AACGAGGTGAACAGCGCGACCGGAGGTGTGATCTCTGAGCCAGACAGATTGTACAG 207
Db      123 ---AspValProTrp-----ValValLeuLeuSerMetAspSerPheTrpLys 137
Qy      208 GTCTCTGACGACAGACAGAAAGCCAGCCTTGAAGACATCAATTTGACCATCCA 267
Db      138 ValLeuThrGlnGlnGlnGlnGlnAlaAlaHisAsnAsnPheAsnPro 157
Qy      268 GATGCTTTGATATGATTTGATGACACAGACTTGAAGACATCGTGAAGGCGAAACG 327
Db      158 AspAlaPheAspPheAspLeuIleIleSerThrLeuLysLysGlnGlyLysSer 177
Qy      328 GTGAGAGTGCGGACCTATGATTTGTGACACACTCAAGCTTACACAGACACAGCTGTC 387
Db      178 ValGlnValProIleTryrAspPheThrHisSerArgLysLysAspTrpLysThrLeu 197
Qy      388 TACCTCTCGGACGCTGTCTGTGTGAGGCGCATCTTGCTTCAACGACGAGATCCGG 447

```

Db 198 TyrclyalaenVallellePhegluglyleMecAlaPheAlaAspLyThrleu 217

Qy 448 GACATGTTCCACCGCGCTCTGTGGACCGACCTCCGACGCTGTCTGAGA 507

Db 218 GluLeuLeuAspMetLysIlePheValAspThrAspSerAspIleArgLeuValArgArg 237

Qy 508 GTTCTCCGGAGCGT---CGCCGAGGAGGAGGACCTGGACAGATTCTGACCGACATACAC 564

Db 238 LeuArgAspAspIleSerGluArgGlyArgAspIleGluGlyValIleLysGlnThrAsn 257

Qy 565 ACCTTCGTGAACCGCGCTTGGAGAGATTCTGCGCCGACAAAGAGTATCCGATGTG 624

Db 258 LysPheValLysProAlaPheAspGlnThrIleGlnProThrMetArgLeuAlaAspIle 277

Qy 625 ATCATCCAGACGAGGTGACCAATATGTTGCCATCCATCCATGATCGTGCAGACATCCAG 684

Db 278 ValValProArgGlySerGlyAsnThrValAlaIleAspLeuIleValGlnHisValHis 297

Qy 685 GACATTCTG 693

Db 298 SerGlnLeu 300

RESULT 8

UCL_CABEL STANDARD; PRT; 229 AA.

ID UCL_CABEL

AC Q17413;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCL) (Uridine monophosphokinase) (Cytidine monophosphokinase).

GN B0001.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Sims M.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.

CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.

CC -1- PATHWAY: Pyrimidine salvage pathway.

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z69634; CA93453.1; -.

DR PIR; T18629; T18629.

DR WormPep; B0001.4; CE05142.

DR InterPro; IPR006082; PRK.

DR InterPro; IPR006083; PRK URK.

DR InterPro; IPR00764; Uridine_kin.

DR Pfam; PF00485; PRK; 1.

DR PRINTS; PRO0478; PHRIBKINASE.

DR PRINTS; PRO0988; URIDINKINASE.

KT Hypothetical protein: Transferase: Kinase; ATP-binding.

FM NP_BIND 15 22 ATP (POTENTIAL).

SO SEQUENCE 229 AA; 26195 MW; 89CB21596A820FCC CRC64;

Alignment Scores:

Pred. No.: 4.03e-31

Score: 512.50

Percent Similarity: 69.05%

Best Local Similarity: 50.00%

Query Match: 33.85%

Length: 229

Matches: 105

Conservative: 40

Mismatches: 58

Indels: 7

DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x UCL_CABEL (1-229)

Qy 67 CCCTTCTGATAGGGGTGACGGCGGACCTCCAGCGGAGTGCACCGTGTGAGAG 126

Db 8 ProLeuLeuIleGlyValAlaGlyGlyThrSerCysGlyLysSerThrIleValGluArg 27

Qy 127 ATCATGAGATGTTCTGGACACAGACGAGTGAACAG---CGGACGGAGAGTGTATC 183

Db 28 IleIleGlnsnuLeuAsnAlaAsnAlaLysGlnSerGlyArgGlnIleAspIleValHis 47

Qy 184 CTGAGCCAGACAGAGTCTCTCAAGATCTTGACGCGACAGACAGAGCGACCTTGAAA 243

Db 48 LeuSerLeuHisSerPheThrArgGluLeuSerAlaGluGluLysIleLeuAlaArgGlu 67

Qy 244 GACAGTACATATTTTGACCATCCAGATGCTTTGATTAATGATTGATGACAGACTTG 303

Db 68 GlyLysPheAsnPheAspHisProAspGlnIleAsnPheAspLeuAlaGluThrLeu 87

Qy 304 AAGAATCTGTGAGAGCGGAAAGCGTGGAGTCCGACCTAATGATTGTGACACTCA 363

Db 88 GlnAsnMetIleAspGlyLysThrValGluIleProLysThrAspMetIleThr---Ser 106

Qy 364 AGGTACACAGACACACGGTGTCTTACCCCTGCGAGCTGTGTTGAGGCACTTG 423

Db 107 SerMetCysnGlyThrValThrValGluProAlaLysValIleIleGluGlyIleLeu 126

Qy 424 GTGTTCTACAGCAGAGATCCGGACAGTTCACCTCGCTTTCTGACACCGAC 483

Db 127 LeuLeuLysAspGluArgValArgLysLeuLeuSerThrLysLeuPheValGluLysAsn 146

Qy 484 TCCGAGCTGACGGTGTCTGCAAGATT-----CTCGGGACGTGGCGGAGGAGG 534

Db 147 AlaGluSerArgLeuArgAsnArgLeuAlaThrTyrlleArgAspThrHisAlaGluPro 166

Qy 535 GACCTGAGACAGATTCTGACGACGATACACCACTTCGTGAAGCGGCTTGGAGATTTC 594

Db 167 -----LeuSerIleIleArgGlnThrGluPheValLysProAlaPheGluGluPhe 184

Qy 595 TGCCTGCCGCAAAAGAAATGTCGATGTGATATCCACAGAGAGTGCACATATGTT 654

Db 185 CysArgProThrLysLysTyrlaAspValIleIleProArgGlyLysAsnHisVal 204

Qy 655 GCCATGAACCTGATCGTGCAGCAGCATCCAG 684

Db 205 AlaThrAspLeuIleAlaLysAsnLeuGln 214

RESULT 9

URK_BACHD STANDARD; PRT; 211 AA.

ID URK_BACHD

AC Q9KDB8;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).

GN URK OR BH1275.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=86665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;

RA "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.

CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.

CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AP001511; BAB04394.1; -.
 DR PIR, C83809; C83809.
 DR HAMAP, MF_00551; -1.
 DR InterPro, IPR003593; AAA_ATPase.
 DR InterPro, IPR006082; PRK.
 DR InterPro, IPR006083; PRK URK.
 DR InterPro, IPR00764; Uridine_kin.
 DR Pfam, PF00485; PRK; 1.
 DR PRINTS, PRO0478; PHRIBLKINASE.
 DR PRINTS, PRO0988; URIDINKINASE.
 DR TIGRPFAM, TIGR00235; udk; 1.
 DR Transfexase, Kinase; ATP-binding, Complete proteome.
 DR KW, NE BIND 12 19 ATP (POTENTIAL).
 DR FT, NE BIND 12 19 ATP (POTENTIAL).
 DR SQ SEQUENCE 211 AA; 24387 MW; C2AFR2CB030520B CRC64;

Alignment Scores:
 Pred. No.: 3,73e-28 Length: 211
 Score: 473.00 Matches: 93
 Percent Similarity: 66.51% Conservative: 50
 Best Local Similarity: 43.26% Mismatches: 56
 Query Match: 31.24% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x URK_BACHD (1-211)

QY 61 CAGCGGCGCTCTCTGATAGGGGTGAGCGGCGACTGCGAGCGGGAAGTGCAGCTGTGT 120
 DB 3 LysArgProIleIleIleGlyValAlaGlyGlyThrGlySerGlyLysThrIleValAla 22
 QY 121 GAGAGATCATGAGTGTCTGCGAGACGAGGTGGAACGAGCGCGAGAGTGTGTC 180
 DB 23 LysGluIlePhe-----TyrGlnPheGlnGlySerIleVal 35
 QY 181 ATCTGAGCCGAGACAGGTTCATCAAG-----GTCTGACGCGAGACGAGAGGCC 231
 DB 36 LeuIleGluIleAspAlaTyrTyrLysAspGlnSerGlnLeuSerLeuGluGlnArgLeu 55
 QY 232 AAGGCGCTGAAAGAGCATATTTTGAACATCCAGATCCCTTGAATATGATTTGATG 291
 DB 56 -----GlnThrAsnTyrAspHisProLeuValaPheAspAsnAspLeuLeu 70
 QY 292 CACAGGACTGTGAAGACATCGTGAAGGCGAAACGGTGAAGGTGCGACCTATGATTTT 351
 DB 71 IleGluHisLeuHisSerLeuLeuAsnGlyGlnAlaIleGluLysProValTyrAspTyr 90
 QY 352 GTACACACTCAAGGTTACAGAGACCAAGCGTGCTTACCTCGCGGACGTGTTCTGTTT 411
 DB 91 LysLeuHisThrArgSerAsnGluValIleLeuValGlnProLysAspValIleIleLeu 110
 QY 412 GAGGCGCATCTGTTGTTCTTACACCGAGAGATCCGGACATGTTCCACCTCGCCCTTTC 471
 DB 111 GlnGlyIleLeuLeuLeuGlnLysArgLeuArgLeuLeuLeuLeuLeuLeuLeu 130
 QY 472 GTGACACCGACTCCGACGTCAAGCTGTCTCGAAGATTCTCCGGGACGTGCC--CGA 528
 DB 131 ValAspThrAspAlaAspIleArgIleIleArgHisMetAlaArgAspIleArgGlnArg 150
 QY 529 GGGAGGAGCTGAGACGATTTCTGACGACGATCAACCACTTGTGAAGCCGCGCTTCGAG 588
 DB 151 GlyArgThrLeuGlnSerValIleGlnGlnTyrThrLysValValArgProMetHisMet 170

QY 589 GAGTCTGCCTGCGGCAAAAGATGCGGATGATCATCCACGAGAGTGCAGCAAT 648
 DB 171 GlnPheIleGluProThrLysArgTyrAlaAspValIleIleProGluIleGlnAsn 190
 QY 649 ATGGTTGCATCAACTGATGTCGACGACATCCAGACATTCGTG 693
 DB 191 ArgValAlaIleAspLeuMetValThrLysIleArgAlaIleIle 205

RESULT 10
 URK_CLOPE
 ID URK_CLOPE STANDARD; PRT; 208 AA.

AC Q8XJ16;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
 DE monophosphokinase).
 GN UDK OR CPE1770.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 ON NCBI_TaxId=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J3 / Type A;
 RX MEDLINE=21664373; PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamaehita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kohara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, AP003191; BAB81476.1; -.
 DR HAMAP, MF_00551; -1.
 DR InterPro, IPR006082; PRK.
 DR InterPro, IPR006083; PRK URK.
 DR InterPro, IPR00764; Uridine_kin.
 DR Pfam, PF00485; PRK; 1.
 DR PRINTS, PRO0478; PHRIBLKINASE.
 DR PRINTS, PRO0988; URIDINKINASE.
 DR TIGRPFAM, TIGR00235; udk; 1.
 DR Transfexase, Kinase; ATP-binding, Complete proteome.
 DR KW, NE BIND 11 18 ATP (POTENTIAL).
 DR FT, NE BIND 11 18 ATP (POTENTIAL).
 DR SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DD5E CRC64;

Alignment Scores:

Pred. No.: 8.84e-28 Length: 208
 Score: 468.00 Matches: 88
 Percent Similarity: 66.51% Conservative: 55
 Best Local Similarity: 40.93% Mismatches: 56
 Query Match: 30.91% Indels: 16
 DB: 1 Gaps: 4

US-09-896-522-3 (1-834) x URK_CLOPE (1-208)

QY 61 CAGCGGCGCTCTCTGATAGGGGTGAGCGGCGACTGCGAGCGGGAAGTGCAGCTGTGT 120
 DB 2 LysArgProIlePheIleIleGlyValAlaGlyGlyThrGlySerGlyLysThrIleValAla 21
 QY 121 GAGAGATCATGAGTGTCTGCGAGACGAGGTGGAACGAGCGCGAGAGTGTGTC 180

Dd		22	LysGullleTyraArgInlPheglYgluaSp-----CysIleAla	34
Oy		161	ATCCGAGGCCAGACAGATTCTTACAAG-----GTCCGACGCAGACAGAAGCC	231
Dd		35	MettIeGluglnAspserTYrTYrLysAsnGlnSerHISLeuSerMetGluAspArGVal	54
Oy		232	AAGGCGCTTAAGAAGACAGTACAATTTTGACCATCCAGATGCCCTTTATAAATGGATTGAA	291
Dd		55	LysThr-----AsnTYrAspHisProAsnAlaPheAspAsnSndLeu	69
Oy		292	CACAGAGACTGTGAAGAATCGGAGGGCAAAAGCGTAGAGGTGCCAGCTATGATTTT	351
Dd		70	ValserrHisLeuGlnSerLeuSndengLYNHSerIIEglnLysProSerTYrAspHe	89
Oy		352	GTGACACACTGAAGTTACACAGAACCAAGCTGGTTCACCTCCGACCTGGTTCGTTT	411
Dd		90	SerIIEHsAsnArGLegIuAspThrThryLysValGlnProlysIuleValIleVal	109
Oy		412	GAGGCGATTGTGTGTCTTACACGCCAGAGATCCGGAGCATGTTCCACTCGGCTCTTC	471
Dd		110	GluGlyIleLeuIleLeuengLnaSPProArgIIeArgGlnLeuSndaspIleLysIleTy	129
Oy		472	GTGACACCGGACCTCGAGCTGAGCGGTCTCGAAGAGTCTCCGGAGAGCTG---CGCGCA	528
Dd		130	ValsAspThrAspAlasppAlarGLieIlleaYGagMeValArGspIleasnGlnUarg	149
Oy		529	GGGAGGACCTTGAGACGATTCGTGACGCAGTACCACACCTTGTAAGCGCGCTTCGAG	588
Dd		150	GLYAArgThnMeGlnSerValIleasnGlnTYrLeuAsnValValLysPromethIsAsn	169
Oy		589	GAGTTCTGCTCCCGACAAAGAATGATGCCGATGTGATCATCTCCAGAGAGTGAACAAT	648
Dd		170	GlnPheThnGlnPProThryLysPheAlaAspIleIleIleProGlnGlnGlyHisAsn	189
Oy		649	ATGGTTCGATCAACCTGATCGTGCAGCACATCCAGAGATCTTG	693
Dd		190	LysValAlaIleAspIleValValaLysIleLysGlnValLeu	204
<hr/>				
RESULT 11				
ID	URK_LACLA	STANDARD;	PRT;	206 AA.
Ac	09CFE21;			
Dt	28-FEB-2003 (Rel. 41, Created)			
Dt	28-FEB-2003 (Rel. 41, Last sequence update)			
Dt	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).			
GN	UDK OR LIL660.			
OC	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.			
ON	NCB1_TaxID=1360;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=IL1403;			
RC	MEDLINE=21235186; PubMed=11337471;			
RA	Bojotin A., Wincker P., Mauger S., Jallou O., Malarne K., Weissenbach J., Ehrlich S.D., Sorokin A.;			
RT	"The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403."			
RL	Genome Res. 11:731-753(2001).			
CC	-1 CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.			
CC	-1 CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.			
CC	-1 PATHWAY: Pyrimidine salvage pathway.			
CC	-1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1 SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.			
CC	-----THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sfb.ch/announce/			

CC		or send an email to license@sb-sib.ch).	
CC	-----		
DR	EMBL; AE006396; AK05758.1; -		
DR	PIR; D86832; D86832.		
DR	HAMAP; MF_00551; -; 1.		
DR	InterPro; IPR006082; PRK.		
DR	InterPro; IPR006083; PRK URK.		
DR	InterPro; IPR00764; Uridine_kin.		
DR	Pfam; PF00485; PRK; 1.		
DR	PRINTS; PRO0478; PHRI.B.KINASE.		
DR	PRINTS; PRO0988; URIDINKINASE.		
DR	TIGRFam; TIGR00235; udk; 1.		
KW	Transferase; kinase; ATP-binding; Complete proteome.		
FT	NP_BIND 11	ATP (POTENTIAL).	
SQ	SEQUENCE 206 AA; 23740 MW; 82E3BF052037DBEE CRC64;		
Alignment Scores:			
Pred. No.:	9,996-27	Length:	206
Score:	454.00	Matches:	92
Percent Similarity:	62.50%	Conservative:	43
Best Local Similarity:	42.59%	Mismatches:	65
Query Match:	29.99%	Indels:	16
DB:	1	Gaps:	5
US-09-896-5522-3 (1-834) x URK_LACLA (1-206)			
OY	61 CAGCGGCCCTTCTGATATAGGGGTGACCGCGCAGCTGCACGCGGAAGTCACGTGTGT	120	
Db	2 Lyslysthrleullellelgllyalrnlrgylgserrallasercllylsrthevalser	21	
OY	121 GAGAAATCATGAGACTGCTCTGGGACAGAACGAGGTGAAACAGCGCACGGAAGGTGCTC	180	
Db	22 HisalalleugluThrPhe---SerAsnGu-----Arglleala	34	
OY	181 ATCTGACCACGAGCAAGTTCTACAG-----GTCTGACGCGCAGACGAGAAGGCC	231	
Db	35 MetlllegluHsaSpSerTyrrtyrlysaspGlnserHslenuThrlphegluIunlrghr	54	
OY	232 AAGCGCTTGAAAGGACAGTAAATTGTGACCATTCAGATGCTTGTGATATGATTTGATG	291	
Db	55 LysThr-----AsntyrAsphSPROLeuAlaPhasPrThrAspyrLreu	69	
OY	292 CACAGACTTGAAAGACATCGTGGAGGGCAAAACGCTGAGAGTCCGACCTGATATTT	351	
Db	70 llealagInleuLygIuLeuglntyrGlYargAlaValaspilleProIIeTryAsprTyx	89	
OY	352 GTGACACACTGAAGTTACCGAGACCAACGCGTGTCTACCTCGCGAAGTGTCTGTTT	411	
Db	90 AlaIyshIshtnrArGserGelngluThrTyrrArglngluProValaspvalleuleval	109	
OY	412 GAGGCACTTGTGTGTCTTACAGCCAGAGAGATCCGAGACATGTTCCACTCGGCTCTTC	471	
Db	110 GluGlYlleleuValleuGluspsoluArgleuarGbaPleumeMetAspIIeIysllePhe	129	
OY	472 GTGACACCGACTCCGACGTCGAGCGTGTCTCGAAGATGTTCTCCGGAAGT---CGCGCA	528	
Db	130 ValAsprTrAspaSpAspValArgIlleIeaArgarGlleaArgAspIIeIePGluGlyalserHn	149	
OY	529 CGGAGGCACTCGAGCAAGATTCGACGAGTAACACCACTTCGTGAAGCGGCGCTTCGAG	588	
Db	150 GlylrgrThrleuAspServallIethrGlnTyrrleuAspAlaValIlyAsproMetTyrrHis	169	
OY	589 GAGTTTGCCCTCGCGCAAAAGATGATCGGATGATGATCATCCCAAGAGAGTGCACAT	648	
Db	170 GinPheIIeGIuProtnHyrsrTyrrlaSpvallIellefPGluGlyalserHn	189	
OY	649 ATGGTTCATCAACCTGATCGTCAGACACATCCAGCATTTTGAAT	696	
Db	190 ThrValGIyValaspIIelleThrTyrrylellelaaserIIeIeuAen	205	
RESULT 12			
ID__URK_STAMM	STANDARD;	PRT,	207 AA.


```

AC Q997N8; 2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine Kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase) .
GN UDK OR SAV1611 OR SAL439 OR MM1561 .
OS Staphylococcus aureus (strain Mu50 / ATCC 700659) ,
OS Staphylococcus aureus (strain N315) , and
OS Staphylococcus aureus (strain MM2) .
OC Bacteria; Firmicutes; Bacillales; Staphylococcus .
NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700659, and N315;
RX MEDLINE=21311952; PubMed=11416146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mitutani U.-I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Katsuhisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratsuku K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus ."
RL Lancet 357:1225-1240(2001) .
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=23040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiratsuku K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA ."
RL Lancet 359:1819-1827(2002) .
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity) .
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch) .
CC -----
CC EMBL; AP003362; BAB57773.1; -
CC EMBL; AP003134; BAB42703.1; -
CC EMBL; AP004827; BAB95426.1; -
CC PIR; B89943; B89943.
CC HAMAP; MF_00551; -; 1.
CC InterPro; IPR006082; PRK.
CC InterPro; IPR006083; PRK URK.
CC InterPro; IPR007654; Uridine_kin.
CC Pfam; PF00485; PRK; 1.
CC PRINTS; PR00478; PHRIBKINASE.
CC PRINTS; PR00988; URIDINKINASE.
CC TIGRFAMs; TIGR00235; udk; 1.
CC Trnsefame; Kinase; ATP-binding; Complete proteome.
FT NP BIND 11
FT SEQUENCE 207 AA; 23505 MW; FD3A6D465531DBD CRC64;

```

US-09-896-522-3 (1-834) x URK_STRAAM (1-207)

```
OY      CTGATAGGGGTGAGCGGGGCGACCTCCAGCGGGAAAGTGCACCCTGTGTGAAGAAGATCATG 132
        ::::::::::::::::::::
Db       6   lletleeglyllealaglylserselyserglylvsythrtthValthransnullemet 25
OY      GAGTTTGCTGGGAGCAGAACGAGGTGMAcAGCGGGCAGCGGAAGTGGTCATCCCTGAGCCAG 192
        ::::|
Db       26   LysAsnLeu-----glutylHisSerValAlaleuleualagln 38
OY      GACAGGTTCTAC-----AAGTCTGACGCGAGACAGACAAGGCCAAGGCAAGCTTGA 243
        ::::|
Db       39   Aspyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 56
OY      GACACAGTCAATTGTCACCATCCAGATGCTCTTGATATGATTTGATGACACAGACTCTG 303
        ::::::::::::::::::::
Db       57   -----AsnTyraPhisIstProheAlaPhePaasAspleuileuileglaenLeu 73
OY      AAGAACATCGTAGAGGGCAAAAASGTGTGAGGGCGGACCTTAATGATTTGTGACACATCA 363
        |||||
Db       74   LysAspLeuLysAsnGilyrAlaValalaValaValaProtnTrtYrAspyrtyrAlaserHistr 93
OY      AGGTTTACAGAGACCCAGSGGTGCTTACCCCTGCGAGCTGTGTTCTGTGTGAGGGCATCTTG 423
        |||
Db       94   ArgSerArpIleThrIleAspRheLysProLysAspValIleIleValaGluglyIlePhe 113
OY      GTGTTTCTACAGCCSAGAGATCCGGGACATGTTCCATCGCTGCGCTCTCTGTCACACCCAG 483
        ::::|
Db       114   AlaleuGluAnLysValleuYrtaRmeMetAsrValIleIleValaIsprThrasp 133
OY      TCCGACGTGACGGCTGTCTGSAAGAATTTCTCCGGAGCGTGCSCG---CGAGSAGAGGCACTG 540
        ::::::::::::::::::::
Db       134   AlaAspLeuAlaGllleuAlaRghLeuThrArgAspThrLysGlnAlaGlyIyrGerMet 153
OY      GAGCAGATTTGACGCGACGTACACCACTTCGTGAAGCCGACCTTGAGAGGTTTCTGCTG 600
        ::::|
Db       154   AspSerValIleIleasnGlnTyrlenseuSerValValaXrpromechIAspRInPheIlegln 173
OY      CCGACAAAGAAAGTATGCGGATGTATCATCCCAACAGAGTGSACATATGTTGGCCATC 660
        |||||
Db       174   ProThrLysLytyrAlaAspIleIleIleProGlnGlyIySerAsnLysValaIaIle 193
OY      AACCTGATCGTGCAGCATCCAGACATTTCTGAAT 696
        ::::::::::::::::::::
Db       194   AspIleMetThrThrLysIleGlnSerLeuValiser 205
```

RESULT_13
URK_STRP8
ID_URK_STRP8 STANDARD, PRT, 208 AA.
AC_OBP0F8:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cydilide monophosphokinase).
GN UDK OR SPYM18.1380.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Filimicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGS8232 / Serotype M18;
RX MEDLIN=21927593 PubMed=11917108;
RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Syiva G.L., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapuri V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).

Alignment Scores:		
Pred. No.:	1,24e-25	Length: 208
Score:	439.50	Matches: 87
Percent Similarity:	63.01%	Conservative: 51
Best Local Similarity:	39.73%	Mismatch: 64
Query Match:	29.03%	Indels: 17
DB:	1	Gaps: 5

US-09-896-522-3 (1-834) x URK_STRP8 (1-208)

QY		61	CAGCGGCTTCGTATAGAGGGGTGCACCGCAGACTGCCAAGCGGAAGTCAACCGTGTT	120
Db		3	LysylsPProilletllelletllleGllyalThrldlglyserglgylglyThrservalSer	22
QY		121	GAGAAGATCATGAGACTTGCTGGGACAAGAACGAGTGGAAACAGCGGACCGAAGTGGTC	180
Db		23	ArgAlatlleuAspSerPhe-----ProAsnalArqIlela	35
QY		181	ATCCGACCCAGACAGAGTTCACAG-----GTCCGACCGGACAGACAGAGCC	231
Db		36	MettlleGlnHlsAspSerTyrrTyllysaspGlnserHshweSerPheGlnGluArgAla	55
QY		232	AAGCGCTTGAAAAGCACAGTACATTTTGACCATTCCGATGCGCTTGATATGATTTGATG	291
Db		56	LysThr-----AsnTyrAspHisProleualAlaPheAspThrAspPheMet	70
QY		292	CACAGGACTCTGGAAGACATCGTGGAGGGCAAACGGTGGAGAGTCCGACCTATGATTTT	351
Db		71	IlegInGlnLeuLysGlnLeuLeuAlaGlyArgProValAspIleProIleTyrAspIyxr	90
QY		352	GTGACACACTCAAGGTTACAGAGCCACCGGCGTCTACCGTGGGAGTGGTCTGTGTT	411
Db		91	LysLysHsrStrHrAspSerAsnThrThrPheArgGlnAspProGlnAspValilleleVal	110
QY		412	GAGGCGATCTTGTTGTGTTCTTACAGCCAGAGAGATCCGGAGCATGTTCCACTCGACTTTC	471
Db		111	GlnGlyIleleuValleuGlnAspGlnArgLeuArgLeuArgAspLeuMetAspIleLysleuPhe	130
QY		472	GTGACACCGACTCCGAGCTGAGGCTGTCTCGAAGATGTTCCGGAG--GTGCGCGCA	528
Db		131	ValAspThrAspAspIleArgIlelleArgArgIleLysArgAspMetMeGluArg	150
QY		529	GGGAGGACCTGGAGACAGATTCTGACGACAGTACACACCTTCGTGAGCCGGCTTCGAG	588
Db		151	GlyTrsSerleuGlnSerIleleAspGlnIlyrThrservalIlyAspMetMetTyrHis	170

RA Ferretti J. J., Mcshan W. M., Ajdic D. J., Savic D. J., Savic G., Lyon K.,
RA Primeaux C., Seate S., Suvorov A. N., Kenton S., Lai H. S., Lin S. P.,
RA Qian Y., Jia H. G., Najjar F. Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S. W., Roe B. A., McLaughlin R.,
RT "Complete genome sequence of an *M. strain* of *Streptococcus pyogenes*." ,
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
NN [2]

CC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "genome sequence of a serotype M3 strain of group A Streptococcus:
RT phase-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Kawazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "the genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1 CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1 PATHWAY: Pyrimidine salvage pathway.
CC -1 SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1 SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AEO06574; AAK34195.1; -
DR EMBL; AEO14517; AAM79649.1; -
DR EMBL; AP005143; BAC63913.1; -
DR HAMAP; MF_005511; - 1.
CC InterPro; IPR006082; PRK.

```

DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23816 MW; 510007D58B379DB6 CRC64;

Alignment Scores:
Pred. No.: 1,24e-25 Length: 208
Score: 439.50 Matches: 87
Percent Similarity: 63.01% Conservative: 51
Best Local Similarity: 39.73% Mismatches: 64
Query Match: 29.03% Indels: 17
DB: Gaps: 5

US-09-896-522-3 (1-834) x URK_STRPY (1-208)

QY 61 CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGCACTCCAGCGGAAGTCGACCGTGT 120
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 3 Lysylserprolletlellellylvalthrglyserglylserglylserlthrserv 22
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 121 GAGAGATCATGAGATTGCTGGAGACAGAGAGGTGAGACGCGGACGGAAGGTGTC 180
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 23 ArgAlalleuaspserphe-----ProAsnAlaArgIleAla 35
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 181 ATCTTGAGCCAGCAGACAGAGTTCTACAG-----GTCTGACGCGAGACAGAGAGCC 231
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 36 MetileglnthaspserlyrlylvalaspqlnserhismetserPheglugluArgVal 55
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 232 AAGCCCTTGAAGAGCAGTACATTTGACCATTCAGATGCTTGTATTAATGATTATG 291
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 56 Lysrthr-----AsnTyrAspHisIleProleuAlaPheAspThrAspPheMet 70
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 292 CACGAGACTGTGAGACATCGTGGAGGAGAAACGGTGGAGGCGGACCTATGATTTT 351
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 71 IleglndlnleuylsgluLeuAlaIleAlaProValAspIleProIleTyrAspTyr 90
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 352 GTGACACACTCAAGATTACAGACAGACAGCGGTGCTACCCCTCGGACGTGTTCTGTT 411
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 91 LysylserhistrArgserAsnThrThrPheArgIlnAspProglndAspValIleleVal 110
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 412 GAGGGCATCTTGGTGTCTACAGCCAGAGATCCGGACATGTTCCACCTCGGCTTTC 471
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 111 GluGlylleuValleuGluAspGluArgLeuArgAspLeuMetAspIleLysleuPhe 130
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 472 GTGACACCGACGCTCGACGCTGAGGCTGTCGAAAGAGTCTCCGGGAC---GTGCGCGGA 528
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 131 ValAspThrAspAspIleArgIleIleArgArgIleLysArgAspMetMetGluArg 150
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 529 GGGAGGACCTGAGACGATTCGACGACAGACCACTTCGTGAAGCGGCGCTTCGAG 588
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 151 GlyArgserleuGlnserIleleAspIlnTyrThrserAlaValLysPromeTyrHis 170
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 589 GAGTTCGCTCGCCGACAAAGAGATAGCCGATGTATCATCCACAGAGAGTGCACAT 648
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 171 GluPheIleGluProserlyArgTyrAlaAspIleValleProglugluValserAsn 190
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 649 ATGGTTCACATCAACGATCGTGCAGACACATCCAGACATTCGTGATGGAGCATC 705
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 191 ValValAlaIleAspValIleAsnserLysIleAlaSerIleLeu---GlyGluVal 208
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 15
URK_STRPN STANDARD; PRT; 212 AA.
AC 097017;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
monophosphokinase).

```

```

GN UDK OR SP1208 OR SPR1090.
OS Streptococcus pneumoniae, and
OC Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=1163916;
RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heideberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwim M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Unayen L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feildlyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Bargett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAben S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Askunas S.R., Rosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6";
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007421; AAK75315.1; -.
DR EMBL; AE008482; AAK99893.1; -.
DR PIR; A99008; A99008.
DR PIR; B95140; B95140.
DR HAMR; SP1208; -.
DR HAMR; ME 00551; -.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 ATP (POTENTIAL).
SQ SEQUENCE 212 AA; 24466 MW; 47B405D377C70A8A CRC64;

Alignment Scores:
Pred. No.: 1.61e-25 Length: 212
Score: 438.00 Matches: 88
Percent Similarity: 60.75% Conservative: 42
Best Local Similarity: 41.12% Mismatches: 68
Query Match: 26.93% Indels: 16

```

DB: 1 Gaps: 4
US-09-896-522-3 (1-834) x URK_STRPN (1-212)

OY 64 CGGCCCTTCGTATAGGGGTGAGCCGCGCACTGCCAGCGGAAAGTCGACCGTGTGTAG 123
|||
:::|
4 ArgProIleIleIleGlyValThrGlySerGlyGlySerValSerArg 23
|||
124 AAGATCATGAGTTCCTGGACAGAACAGAGTGAACAGCGGACGGAAGGTGTCATC 183
|||
24 AlaIleLeuSer-----HisPheProAspGluValIleSerMet 36
|||
184 CTGAGCCAGACAGGTTCTACAG-----GTCTGACGCGCAGACAGAGGCCAAG 234
:::|
37 IleGluHisAspSerTyrTyrIleAspGlnSerHisLeuThrPheGluGluValValys 56
|||
235 GCCTTGAAAGACATACATTTTGAACATTCAGATGCTTGTGATTAATGATTTGATGAC 294
|||
57 Thr-----AsnTyrAspHisProPheAlaPheAspThrAspLeuMetIle 71
|||
295 AGGACTCTGAAGAACATGTCGAGGCAAAACGCTGAGAGTGCAGACCTATGATTTGTG 354
:::|
72 GluGlnIleIleGlyLeuLeuAlaGlyArgProValAspIleProThrTyrAspTyrThr 91
|||
355 ACACACTCAAGGTTACACAGAGACACGGTGTCTACCTCGCGACGTGGTTCTGTTTGAG 414
|||
92 GluHisThrArgSerSerTyrTyrArgGlnGluProGlnAspValPheIleValGlu 111
|||
415 GGCATCTTGAGTGTCTACAGCCAGACAGATCCGCGGACATGTCACCTGCGCTCTTCGTG 474
|||
112 GlyIleLeuValLeuGluAspIleArgLeuArgAspLeuMetAspIleIleVal 131
|||
475 GACACCGACTCCGACGTCAGCTGTCTGAAAGATTCTCGGAGACGTG---CGCCGAGGG 531
|||
132 AspThrAspAspAspValArgIleIleArgArgIleIleArgAspMetGluGluArgGly 151
|||
OY 532 AGGACCTGAGACGATTCTGACGACGTAACACACCTTCGTGAACCGCGCTTCGAGGAG 591
|||
152 ArgSerLeuAspSerValIleAsnGlnTyrLeuGlyValValysPrometTyrHisGln 171
|||
OY 592 TTCGCGCTCGGACAAAGAGATGCGATGATGATCCACGAGAGAGTGAACATATATG 651
|||
172 PheIleGluSerThrIleArgTyrAlaAspIleValIleProGluGlyValSerAsnThr 191
|||
OY 652 GTTGCCATCAACTGATCGTCAGACATCCAGACATTTCTG 693
|||
192 ValAlaIleAspLeuLeuThrThrIleIleAlaIleValSer 205
|||
Db

Search completed: November 25, 2003, 07:32:11
Job time : 18.0094 secs

Db 82 GluLysIleValGluArgLeuGly-----IleProTrpValThr 94

Oy	18	ATCTGAGCCAGGACGAGTTTCAACAAGCTCTGACGGACGACGAGGCAAGGCAAGCCCTTG	240
Db	95	IIleuSerMetcAspSerPheTyrIysValIleuThrProGluGluIleIysAlaHis	114
Oy	241	AAAGCAGTACAAATTTTGCATCCATCCAGATGCTTTGATTAATGATTGTAGCCACAGGACT	300
Db	115	GIuSerArgTyrAsnPheAspGlyProAlaAlaPheAspPheAspLeuTyrGluVal	134
Oy	301	CTGAGAAACATCGTAGAGGCAAAACCGGTGAGAGGCCGACCTATGATTTGTGACAC	360
Db	135	LeuIysArgLeuArgGluIlyIysSerValAspValProValTyrAspPheAsnThrHis	154
Oy	361	TCAAGTTTACACAGACCAACGGTGTCTACCCCTGCGACGTGGTTCTGTTGAGGGATC	420
Db	155	SeArgAspProAsnSerIysMetMetTyrGlyAlaAspValIleuIlePheGluGlyIle	174
Oy	421	TTCGTGTTTCAACGACGAGAGATCCGGGACATGTTCCACCGCGGCTCTTCGTGACACC	480
Db	175	LeuAlaPheHisAspGluIArgIleuIlyIysValAsnLeuMetAspMetLysValAlaPheValAspThr	194
Oy	481	GACTCCGACGCTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCCG--CGAGGAGGAGAC	537
Db	195	AspGlyAspPheuArgLeuAlaArgArgIleValAlaArgAspValThrAspArgGlyIArgAsp	214
Oy	538	CTGAGACAGATTCTGACGACATACACCAACTTCGTGAAGCCGGGCTTCGAGAGATTCTGC	597
Db	215	IIleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle	234
Oy	598	CTGCGCAAAAGAGTAGTCCGATGTGATCATCCACAGAGAGTGGACAATATGTTGCC	657
Db	235	AlaProCysMetcAspSerAlaAspLeuIleValProArgGlyGlyGluAlaAspValAla	254
Oy	658	ATCAACTGATCTGCGACGACATC	681
Db	255	IIleAspMetIleValGlnAsnVal	262

```

RESULT 2
T21110
hypothetical protein F19B6.1b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21110
R:Thomas, K.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19375
A:Accession: T21110
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-555 <M11>
A:Cross-references: EMBL:Z69635, NID:g1200023; PIDN:CAA93462.1; GSPDB:GN00022; CESP:F19B6
A:Experimental source: clone F19B6
A:Gene: CESP:F19B6.1b
A:Map position: 4
A:introns: 35/1; 83/3; 201/3; 267/3; 395/2; 511/3

Alignment Scores:
Pred. No.:          2,41e-36          Length:          555
Score:              559.00            Matches:          110
Percent Similarity: 74.04%            Conservative:     44
Best Local Similarity: 52.88%          Mismatches:      46
Query Match:        36.92%            Indels:          8
DB:                  2                Gaps:            2

US-09-896-522-3 (1-834) x T21110 (1-555)

OY      61  CAGCGGCCCTTCTCGATAGGGGTGAAGCGGCGGCACTGCCACGCGGAAGTCCAGCCGTGTGT 120
      :: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      102  LysHSPProPheValIIeGIyValCysGIyGSerAlaSerGIyLysThrValAla 121

OY      121  GAGAAAGATCAATGAGACTTGCTGGGACAGAAACGAGTGGGAACAGCGGACGGAAGTGTGTC 180
      |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      122  GltuIySrlIeValIGluArgLeuGIy-----IleProTrrValThr 134

```

QY	181	ATCCGAGCCAGGAAGGTTCTTACAGAGGCTCCGAGCGGACAGGACGAGGCGCAAGGCTTGC	240
Db	135	IleuSerMetAapSerPheTyrIleValIeuThrProGluGluIleValaIalaIle	154
QY	241	AAAGGACAGTCAATTTTGACCATCCAGATGCGCTTTGATATGATTTGATGCACAGGACT	300
Db	155	GlusSerArgTyrAsnPheAspGlyProAsnAlaPheAspPheAspLeuTyrGluVal	174
QY	301	CTGAAGAACAATCGTGGAGGGCAAAACGGGAGGAGCGCACTATGATTTTGACACAC	360
Db	175	LeuIysaArgLeuAlaArgGluGlyLysSerValaIaPvalProValTyrAspPheSerThrHis	194
QY	361	TCAAGTTTACAGAGACCAACGGTGGTCTAACCTCGCGGACGTGGTTCCTGTTGAGGCGATC	420
Db	195	SerArgAapProAsnSerLysMetMetCylrGlyAlaAspValIleuIlePheGluGlyIle	214
QY	421	TTGGTGTTCTTCACGCGACGATCCGGGACATGTTCCACCTCGGCGCTCTTCGGGACACC	480
Db	215	LeuAlaPheHisAspGluAlaGlyIleLysaSerLysMetAspMetLysValaPheValaPheIrr	234
QY	481	GACTCCGACGTCAGGCTGTCTTCGAAGAAGTTCGCCGAGCTGGCGC--CGAGCGAGGAC	537
Db	235	AspGlyAspLeuArgLeuAlaArgArgIleValaIArgAspValThrAspArgGlyArgAap	254
QY	538	CTGAGACGAGATTCTGACCGCATGACACCACTTCGGAAGCGCGCTTCGAGGAGATTCTGC	597
Db	255	IleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLysTyrIle	274
QY	598	CTGCCGACAAAGAAAGTATGCGGATGTGATTCATCCACGAGAGTGGACATATGTTGCG	657
Db	275	AlaProCysMetAspSerAlaAspLeuIleValProArgGlyGlyGluLysAspValaIa	294
QY	658	ATCAACCTGATCTGTCCAGCACATC	681
Db	295	IleAspMetIleValGlnAsnVal	302

```

RESULT 3
T18629
hypothetical protein B0001.4 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C|Accession: T18629
R:Sims, M.
submitted to the EMBL Data Library, February 1996
A|Reference number: Z18995
A|Accession: T18629
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-229 <EMBL>
A|Cross-references: XREF:Z69634; PIDN:CAA93453.1; GSPDB:GN00022; CESP:B0001.4
A|Experimental source: clone B0001
C|Genetics:
A|Gene: CESP:B0001.4
A|Map position: 4
A|Introns: 21/3; 41/3; 108/2
C|Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 1.03e-32 Length: 229
Score: 512.50 Matches: 105
Percent Similarity: 69.05% Conservative: 40
Best Local Similarity: 50.00% Mismatches: 58
Query Match: 33.85% Indels: 7
DB: Gaps: 4

US-09-896-522-3 (1-834) x T18629 (1-229)
QY 67 CCCTTCGATAGGGGTGAGCGCGGCACTGCAGCGGAAGTCGACCTGTGTGAAG 126
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 8 ProlenleuileglyAlaIdyelythrseryCysglylybserthrlvalGluang 27
127 ATCATGAGTGTGCTGGGACAGAAACGAGGTGTAACAG--CGCAGCGGAAGGTGTCATC 183

```

[illegible]

```

RESULT 4
C83809
uridine kinase udk [imported] - Bacillus halodurans (strain C-125)
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: C83809
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliophilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: C83809
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA04994.1; GSFDB:GN00
C/Experimental source: strain C-125
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 1.38e-29 Length: 211
Score: 473.00 Matches: 93
Percent Similarity: 66.51% Conservative: 50
Best Local Similarity: 43.26% Mismatches: 56
Query Match: 11.24% Indels: 16
DB: Gaps: 4

US-09-896-522-3 (1-834) x C83809 (1-211)

```

[illegible]

```

RESULT 5
T41020
probable Uridine kinase - fission yeast (Schizosaccharomyces pombe)
C|Species: Schizosaccharomyces pombe
C|Date: 03-Dec-1999 #sequence_rev100 03-Dec-1999 #text_change 03-Dec-1999
C|Accession: T41020
R|Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A|Accession: T41020
A|Reference number: Z21965
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-454 <MOR>
A|Cross-references: EMBL:AL023860; PIDN:CAA19591.1; GSPDB:GN00068; SPDB:SPCC162.11c
A|Experimental source: strain 972h-; cosmid c162
C|Genetics:
A|Gene: SPDB:SPCC162.11c
A|Map position: 3

Alignment Scores:
Pred. No.:      2 02e-28      Length:      454
Score:          459.00      Matches:      92
Percent Similarity: 62.33%      Conservative: 47
Best Local Similarity: 41.26%      Mismatches:  76
Query Match:    30.32%      Indels:      8
DB:             2          Gaps:         2

US-09-896-522-3 (1-834) x T41020 (1-454)

```

```

Oy      61 CAGCGGCCCTTCTTGATAGGGGTGAACGGCGGCACATGCCAGCGGGAAGTCGACCGTGTG 120
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      3 LysArgProIleIleIleGlyValAlaGlyIleThrlYseGlyLysThrThrValAla 22

```

QY 31 AGCCCCGCGCGAGGCGACCGTCCGCACCAAGCGGCCCTTCTGATAGGGGTAGCGGC 90
||| ||| ::||| ::::: |||||:|||||
Db 9 SerSerAsnProThrTyrglnProProTyrPargLysValArgPheIleGlyIlealagly 28

```

QY 91 GGACCTCCACGGGAGATGACCGTGTGAGAGATCATGGAGTTCTGGGACAGAAC 150
Db 29 ProserglyserglythrnservalalglleuilevallysalaleuanleuPro 48
QY 151 GAGGTGAMACAGCGGACGGAGAGGATGATCTCTGAGCCAGACAGTTCTACAGGTC 210
Db 49 His-----ValValileuenserleuenserperthyrylyser 61
QY 211 CTGACGGCAGACGAGAGCCAGAGGCTTGAAAGACATTCATTTTGACATCCAGAT 270
Db 62 LeuasnlaagluglnlylysalalpheasnlanparyrphaspheasrperProglu 81
QY 271 GCCTTATATATGATTGATGACACAGACTCTGAAGACATCGTGAAGGCGAAACGGTG 330
Db 82 Alalleasprtrphaleuileuphevallyleuileuuglysginglyrlygylval 101
QY 331 GAGGTGCGACCTTATGATTTTGTGACACACTCAAGGTTACAGACACAGGTGTCTAC 390
Db 102 AsplleprolietysrPheasnlglnhlsasnargneuprogliuthranthrleuPhe 121
QY 391 CTGCGGACGCTGCTTCTTTGAGGACATCTGTGTTTCACGCCAGAGATCCGGAC 450
Db 122 GlylaaserllellelleuuglyllephealaleutyraapgluylleargSer 141
QY 451 ATGTTTCACCTGCTCTTGTGACACCGACTCCGACGCTGCTGTCTGAGAGTT 510
Db 142 LeuenuaspyalservalpheuasnprthrasperaspyalCysleuserthargyleu 161
QY 511 CTCGGGACGTCGCGC---CGAGGAGGAGCCTGTGAGACATTTCTGACGACATACACC 567
Db 162 AsnrygasprileanthyraarglylArgaspriValleuglylueuglnlytyraAsnly 181
QY 568 TTGCTGAACCGGCTTCGAGGAGTTCTGCTGCCGACAAAGATTCGCGAGTATC 627
Db 182 PhevallyspaserlyrGluasnphavalarghlnleuserTythrAspleulle 201
QY 628 ATCCACGAGAGGACATATGTTGCCATCACTGATCGACGACATCCAGAC 687
Db 202 ValProargglyarghaphasnlyleuAlalleasphetValilleasnphelleargArg 221
QY 688 ATTCTGAAT 696
Db 222 Thrleuser 224

RESULT 6
Db8632
uridine kinase (EC 2.7.1.48) [imported] - Lactococcus lactis subsp. lactis (strain IL140
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86832
R/Bolotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malame, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: D86832
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-206 <STO>
A/Cross-references: GB:AE005176; PID:g12724672; PIDN:AAK05758.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase
C/Keywords: phosphotransferase

```

```

Alignment Scores:
Pred. No.: 4,43e-28 Length: 206
Score: 454.00 Matches: 92
Percent Similarity: 62.50% Conservative: 43
Best Local Similarity: 42.59% Mismatches: 65
Query Match: 29.99% Indels: 16
Gaps: 5

```

```

US-09-896-522-3 (1-834) x D86832 (1-206)
QY 61 CAGCGGCCCTTCTGATTAGGGGTGAGCGCGGACCTCCAGCGGAAATCGACCGTGT 120
Db 2 LyslysthrleullellelleuglyalthrlyglyserlasercolylserValSer 21
QY 121 GAGACATATGAGATGCTGGGACAGACAGAGTGAACAGCGGACGGAAGTGTGTC 180
Db 22 HisAlalleuuglnthrhe---Seranglu-----ArgileAla 34
QY 181 ATCTGAGCCAGACAGGTTCTTACAG-----GTCTGACGCGACAGACAAAGCC 231
Db 35 MetlleuglnhlsasrperlytyrlysaerGlnserHlleuThrpheluglnargthr 54
QY 232 AAGCGCTTGAAGAGACAGTACATTTGACSCATCCGAGCGCTTGATATGATTGAG 291
Db 55 LysThr-----AsnyraapRhsprleuAlapheasprthraspyrlyeu 69
QY 292 CACAGACTCTGMAAGAACATCGTGAAGGCGAAACCGTGAAGTCCGACTATGATTTT 351
Db 70 IleAlglnleuylseugleuuglnlyrlyargAlaValaspleprolietlyraaplyr 89
QY 352 GTGACACACTCAAGTTACCAAGACACACGCTGTCTACCTGCGGACGTGTTCTTT 411
Db 90 AlalyhlsrthrargserglnlythrtyrArgglnlprovalaspyalleulleval 109
QY 412 GAGGGATCTGTGTCTTTCACGCCAGAGATCCGGAGATGTTCCACTGGGCGCTTC 471
Db 110 GlnGlylleuvalleuuglylueuglylueuglylueuglylueuglylueugly 129
QY 472 GTGACACGAGCTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGTG---CGCG 528
Db 130 ValasprthraspyalargllelleargArglleargArglleuglylueugly 149
QY 529 GGGAGGACCTGAGACATTTCTGACGACATACCACTTGTGAGCGCGGCTTCGAG 588
Db 150 GlyArgThrleuasnperAlallethrGlnlyleuasnAlaVallysePrometyrHis 169
QY 589 GAATTCGCTGCGCGACAAAGATGCGGATGATCATCCACGAGGAGGACAT 648
Db 170 GlnpheilleuglyprothrlysaarglyralaspyalilleProdluglyalSeran 189
QY 649 ATGTTGCATCACTGATCGTGCAGACATCCAGACATTTCTGAAT 696
Db 190 ThrValGlyValasplellellethrThrlylleAlaserlleuasn 205

RESULT 7
Db8943
uridine kinase [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: B89943
R/Kuroda, M.; Ohba, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hatcori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:1143146
A/Accession: B89943
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-207 <KUR>
A/Cross-references: GB:BA000018; PID:g13701409; PIDN:BA842703.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

```

```

Alignment Scores:
Pred. No.: 1.59e-27 Length: 207
Score: 447.00 Matches: 88
Percent Similarity: 65.09% Conservative: 50
Best Local Similarity: 41.51% Mismatches: 58

```


Query Match: 29.52% Indels: 16
 DB: 2 Gaps: 4
 US-09-896-522-3 (1-834) x B95140 (1-207)

QY 73 CTGATAGAGGTGAGCGGCGGCACTGCGAGCGGGAAGTGAACCGTGTGTGAGAAATCATG 132
 Db 6 IIEIIIEGIYIEIAAGIYISERGIIYSEGIYSETHIRHVRVAlThranSgIIIEmet 25
 QY 133 GAGTGTGTGGACAGAAACGAGGTGGAACAGCGGCAAGCGGAAGTGTGATCTTGAGCCAG 192
 Db 26 LysAsnLeu-----GluGIYHISerValAlaLeuLeuAlaGln 38
 QY 193 GACAGGTCTTAC-----AAGTCTGTGACGAGCAAGAAAGCCAGGCTTGA 243
 Db 39 AsprTYrTYrTYrLysAspGlnLysHISLeuThrPheAspGlnAspGluGluThr----- 56
 QY 244 GACAGTACATTTTGAACATCCAGATGCTTTGATATGATATTGATGACAGGACTCTG 303
 Db 57 -----AsnTYrAspHISProPheAlaPheAspHISnLeuLeuIIeGIuAsnLeu 73
 QY 304 AAGAACTCTGGAGGCAAAACGTTGAGGTGCGGACCTATGATTTTGTGACACTCA 363
 Db 74 LysAspLeuLysAsnGIYsAlaValGIuValProThrTYrAspTYrAlaSerHISThr 93
 QY 364 AGGTACACAGAGCAACGATGCTTACCTGCGGAGCTGTCTTCTGTGACACCCAG 423
 Db 94 ArgSerAspIIEHrIIEAspPheLysProLysAspValIIEIIEValGIuGIYIIEPhe 113
 QY 424 GTGTTCACAGCCAGAGATCCGGGACATGTTCACCTGTCTTCTGTGACACCCAG 483
 Db 114 AlAlenGIuLysValIIEAspPheMetCAspAlaIIEYsIIEYrValaIIEAspThrAsp 133
 QY 484 TCCGACGTGAGCTGTCTTCCAGAGTTCTCCGGAGCTGGC--CGAGGAGGACCTG 540
 Db 134 AlAspPheAspArgIIEAspArgIIEAspArgIIEAspArgIIEAspArgIIEAsp 153
 QY 541 GAGAGATCTGACGAGTACACACTTGTGAGCGGAGCTTCCAGAGGTTCTGCTG 600
 Db 154 AspSerValIIEAsnGlnTYrLeuSerValValaIIEArgPheMetHISAspGlnPheIIEGlu 173
 QY 601 CCGACAAAGATATGCGGATGTGATCATCCACGAGAGTGTGACATATGTTGCCATC 660
 Db 174 ProThrLysLysTYrAlaAspIIEIIEIIEIIEProGIuIIEYsAsnLysValaIIE 193
 QY 661 AACCTGATCGTGCAGACATCCAGGACATTTCTGAT 696
 Db 194 AspIIEAspLeuThrTYrLysIIEGlnSerLeuValSer 205

RESULT 8
 B95140
 uridine kinase [imported] - Streptococcus pneumoniae (strain TIGR4)
 C/Species: Streptococcus pneumoniae
 C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C/Accession: B95140
 R/Tetelin, H.; Nelson, K.B.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umkam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,
 nson, T.; Hickey, E.R.; Holt, I.E.
 Science 293, 498-506, 2001
 A/Authors: Lotz, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A./Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A/Reference number: A95000; MUID:21357209; PMID:11463916
 A/Accession: B95140
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-212 <KUR>
 A/Cross-references: GB:AE005672; PIDN:AAK75315.1; PID:g1972689; GSPDB:GN00164; TIGR:SP4
 A/Experimental source: strain TIGR4
 C/Genetics:
 A/Gene: SPI208
 C/Superfamily: uridine kinase
 Alignment Scores:

Pred. No.: 8.28e-27 Length: 212
 Score: 438.00 Matches: 88
 Percent Similarity: 60.75% Conservative: 42
 Best Local Similarity: 41.12% Mismatches: 68
 Query Match: 28.93% Indels: 16
 DB: 2 Gaps: 4
 US-09-896-522-3 (1-834) x B95140 (1-212)

QY 64 CGGCGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAGCGGGAAGTGAACCGTGTGTGAG 123
 Db 4 ArgProIIEIIEIIEGIYValThrGIYsERGIIYGIYLYrHISerValSerArg 23
 QY 124 AAGATGATAGTGTGTGGGACAGAAACGAGTGTGAACAGCGGCAAGGAGTGTATC 183
 Db 24 AlaIIEuSer-----HISProAspGlnLysIIESerMet 36
 QY 184 CTGAGCCAGACAGTGTCTTACAG-----GTCTGACGCGGACAGCAAGAGCCAG 234
 Db 37 IIEGIuHISAspSerTYrTYrLysAspGlnSerHISLeuThrPheGIuGIuArgValLys 56
 QY 235 GCGTTAAAGACAGTACATTTTGCATCCAGATGCTTTGATATGATTTGATGCAC 294
 Db 57 Thr-----AsnTYrAspHISProPheAlaPheAspThrAspLeuMetIIE 71
 QY 295 AGGACTGTGAAGACATCGGAGGAGGCAAAACGTTGAGGTGCGGACCTATGATTTTGTG 354
 Db 72 GluGIuIIEYsGIuLeuLeuAlaGIYArgProValAspIIEProThrTYrAspTYrThr 91
 QY 355 ACACTCAAGTGTACAGACACACGATGTCTTACCTGTGCGGACGTGTCTGTGTGAG 414
 Db 92 GluHISThrArgSerSerTYrTYrArgGlnGIuProGIuAspValPheIIEValGIu 111
 QY 415 GCGATCTTGTGTGTCTTACAGCCAGAGATCCGGGACATGTTCCACTGCGCTTCTGTG 474
 Db 112 GIYIIELeuValIIEuGIuLysPysArgIIEAspPheMetCAspIIEYsIIEPheVal 131
 QY 475 GACACGAGCTCGGACGTCGAGGCTGTGAGAGGTTCTCCGGAGCTG--GCGAGAGG 531
 Db 132 AspThrAspAspAspAlaArgIIEIIEIIEArgGIIIEYsArgAspPheMetCAspIIEYsIIEPheVal 151
 QY 532 AGGAGCTGTGAGCAGATTTGTGACGAGTACACACTTGTGTAAGCGGCTTCCAGAG 591
 Db 152 ArgSerLeuAspSerValIIEAsnGlnTYrLeuGIYValValLysProMetTYrHISGln 171
 QY 592 TTCTGCTTCCGCAAAAGATATGCGGATGTGATCATCCACGAGAGTGTGACAAATG 651
 Db 172 PheIIEGIuSerThrTYrAspArgTYrAlaAspIIEValIIEProGIuIIEYsAsnThr 191
 QY 652 GTTGCATCAACTGATCTGTCAGACGACATCCAGGACATTTCTG 693
 Db 192 ValaIIEAspLeuLeuThrTYrLysIIEAlaLysIIELeu 205

RESULT 9
 A99008
 uridine kinase (EC 2.7.1.48) [imported] - Streptococcus pneumoniae (strain R6)
 C/Species: Streptococcus pneumoniae
 C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C/Accession: A99008
 R/Hoskins, J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Buzget, S.; Dehoff, B.S.; E
 r, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.U.; Lu, J.; Matsushima, P.; McHenry, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 185, 5709-5717, 2001
 A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A./Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A/Reference number: A97872; MUID:21429245; PMID:11544234
 A/Accession: A99008
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-212 <KUR>
 A/Cross-references: GB:AE007317; PIDN:AAK9893.1; PID:g15458714; GSPDB:GN00174
 C/Genetics:
 A/Gene: udk

Db 185 ValAspAlaAspLeuAspValCysLeuAlaArgArgLeuSerArgAspAlaValSerArg 204
QY 529 GGGAGGGAGCTGGAGGAGTTCGAGCGAGTACCACTTCGGAAGCGGCGCTTCGAG 588
Db 205 G1AArgAspLeuAspGlyCysIleGlnIleProGluLysAspValLysProAsnAlaVal 224
QY 589 GAGTTTCGCTGCGCCGCAAAAGATATGCCAGATGTCATCCACGAGAGTGACAAAT 648
Db 225 LysPheValLysProThrMetLysAsnAlaAspAlaIleIleProSerMetSerAspAsn 244
QY 649 ATGGTTCCATCAACCTGATCGTCGACACATC-----CAGCATTCGGAAGT 699
Db 245 AlAThrAlaValAsnLeuIleIleAsnHisIleLysSerLysLeuGluLeuLysSerAsn 264
QY 682 -----CAGCATTCGGAAGT 699
Db 265 GluHisLeuArgGluLeuIleLysLeuGlySerSerProSerGlnAspValLeuAsnArg 284
QY 700 GACATC 705
Db 285 AsnIle 286
RESULT 11
G69728
uridine kinase udk - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C/Accession: G69728
R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berton
C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.E.
Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois,
A./Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A./Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Toesto, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A./Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A./Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A./Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69728
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-211 <KUN>
A/Cross-references: GB:299117; GB:AL009126; NID:g2634966; PIDN:CAB14675.1; PID:g2635179
A/Experimental source: strain 168
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase
Alignment Scores:
Pred. No.: 3 57e-26 Length: 211
Score: 430.00 Matches: 84
Percent Similarity: 63.26% Conservative: 52
Best Local Similarity: 39.07% Mismatches: 63
Query Match: 28.40% Indels: 16
Gaps: 4
US-09-896-522-3 (1-834) x G69728 (1-211)
QY 61 CAGCGGCGCTTCGTATAGGGGTGAGCGGCGGCACTGCGAGCGGAAGTGCACCGTGTGT 120
Db 3 LysAsnProValValIleGlyIleAlaGlyLysSerLysSerLysValThr 22
QY 121 GAGAAGTATGAGAGTTCGTCGAGACAGAGTGAACAAGCGGCGGCGGAAGTGTGTC 180
Db 23 ArgSerLysIleArg-----GlnPheLysGlyHisSerIleLeu 35
QY 181 ATCTTAGCGCAGACAGAGTTCACAG-----GTCTGACGGCGAGCAGGAAGGCC 231

Db 36 MetIleGlnGlnAspLeuTyrTyrLysAspGlnSerHisIleLeuProPheGluGluArgLeu 55
QY 232 AAGCGCTTGAAAGACAGTACAAATTTTGACCATCCGAGATGCCCTTTATATGATTTGAAG 291
Db 56 AsnThr-----AsnTyrAspHisProLeuAlaPheAspAsnAspTyrLeu 70
QY 292 CACAGAGCTCTGAGAAACATCGTGAGGGCGCAAAAGCGTGAGAGTGCAGCCTATGATTTT 351
Db 71 IleGluHisIleGlnAspLeuAsnTyrArgProIleGluLysProIleTyrAspTyr 90
QY 352 GTGACACACTCAAGTTTACAGACAGACAGCGGTGTCTACCTCGCGAGCGGTTCGTTT 411
Db 91 LysLeuHisThrArgSerGluGluThrValHisValGluProLysAspValIleIleLeu 110
QY 412 GAGCGCATTCGTTGTTCTTACACCGACAGATCCGAGCATGTTCCACCTCGCGCTTTC 471
Db 111 GluGlyIleLeuValLeuGluAspLysArgLeuArgAspLeuSerAspIleLysLeuTyr 130
QY 472 GTGACACAGCGACCGCGAGCGGTGTCTCGAAGGTTCTCGGAGCGG---CGCCGA 528
Db 131 ValAspThrAspAlaAspLeuArgIleIleArgArgIleMetArgAspIleAsnGluArg 150
QY 529 GGGAGGAGCTGGAGCGAGTTCGAGCGCATACACCACTTCGTCGGAAGCGGCTTCGAG 588
Db 151 GlyArgSerIleAspSerValIleGluGlnTyrValSerValValArgPromethisAsn 170
QY 589 GAGTTTCGCTGCGCCGCAAAAGATATGCCAGATGTCATCCACGAGAGTGACAAAT 648
Db 171 GlnPheValLysProThrLysArgTyrAlaAspIleIleIleProGluGlyGlnAsn 190
QY 649 ATGGTTCCATCAACCTGATCGTCGACACATC-----CAGCATTCGGAAGT 699
Db 191 HisValAlaIleAspLeuMetValThrLysIleGlnThrIleLeu 205
RESULT 12
F84277
uridine kinase (imported) sp. Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C/Accession: F84277
R/Ky, W.V.; Kennedy, S.P.; Malafas, G.G.; Berguist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leitbauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freltas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.U.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Lie
A./Title: Genome sequence of Halobacterium species NRC-1.
A./Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: F84277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-248 <STO>
A/Cross-references: GB:AE004437; NID:gl0580750; PIDN:AGI19586.1; GSPDB:GN00138
C/Genetics:
A/Gene: urk
C/Superfamily: uridine kinase
Alignment Scores:
Pred. No.: 3.94e-25 Length: 248
Score: 417.00 Matches: 83
Percent Similarity: 63.64% Conservative: 50
Best Local Similarity: 39.71% Mismatches: 58
Query Match: 27.54% Indels: 18
Gaps: 5
US-09-896-522-3 (1-834) x F84277 (1-248)
QY 70 TTCCTATAGGGGTGAGCGGCGGCACTGCGAGCGGAAGTGCACCGTGTGAGAGATC 129
Db 29 PheAlaIleGlyIleAlaGlyLysThrGlyAlaGlyLysThrValAlaArgGluIle 48
QY 130 ATGAGTTCGTCGAGACAGAGTGAACAAGCGGCGGCAAGAGTGTCTCATCCGAGC 189

Db 49 ThrAspAnValGlyIuSer-----AlaThrLeuIlePro 60
Qy 190 CAGAGAGGTTCTCAAGGTCCTGACG-----GCAGAGCAGAAAGGCCAAGGCTTG 240
Db 61 LeuAspAnTyTyTyGluAspLeuSerAspArgProPheGluGluArgAlaAsnAla--- 79
Qy 241 AAGAGACATGCAATTTTGAACATCCAGATGCTTTGATATATATTTGATGACAGAGACT 300
Db 80 -----AanTyAspHisProSerAlaPheGluTyrGluLeuLeuArgThnHis 95
Qy 301 CTGAAGAACATCGTGGAGGGCAAAACGATGAGGTGCCGACCTTATGATTTGTACACAC 360
Db 96 IleAspAlaLeuLeuSerGlyGlnSerIleGluMetProGlnTyAspPheGluArgHis 115
Qy 361 TCAAGGTTACCGAGACACCGAGGTGTACCTCGGAGCGTGTCTGTTGAGGGATC 420
Db 116 ValArgLysAlaAspArgValValValGluProThrAspValIleValLeuGluGlyIle 135
Qy 421 TTGGTGTCTTACACCGAGAGATCCGGACATGTTCCACCTGCGCTCTTCGTGAGACAC 480
Db 136 LeuAlaLeuSerAspArgIuThrValAlaAspMetLeuAspLeuHisIleTyValGluThr 155
Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGACTTCCGGGAC---GTCCGCGAGAGGAGAC 537
Db 156 AspAlaAspValArgIleLeuArgGlyIleGluArgAspValValGluArgGlyArgGlu 175
Qy 538 CTGAGAGAGATTCGAGCGAGATACACCACTTCGTGAAGCGCGCTTCGAGGATTTGCG 597
Db 176 LeuGluGlyValMetAspGlnTyLeuSerThrValIlyAspPheMetHisGluGlnPheIle 195
Qy 598 CTGCGCAAGAAAGATGCGCATGTGATCATCCACAGAGATGACAAATATGTTGCC 657
Db 196 GluProThrArgLysArgHisAlaAspIleIleIleProGluGlyAla---AanSerValAla 214
Qy 658 ATCAACCTGATCGTCGACGACATCCAG 684
Db 215 ValAsnLeuLeuGluGluValGln 223

RESULT 13
D86582
uridine kinase [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: D86582
R/Shirai, M.; Hixkawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: D86582
A/Molecule type: DNA
A/Status: preliminary
A/Experimental source: strain J138
A/Cross-references: GB:BA000008; NID:98979107; PIDN:BA98942.1; GSPDB:GN00142
A/Genetics:
C/Genetics:
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 9.65e-25 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 27.21% Indels: 22
DB: 2 Gaps: 6

US-09-896-522-3 (1-834) x D86582 (1-222)

Qy 73 CTGATAGGTTGAGCGGAGCGGACGCAAGTCGACCGTGTCGAGAGATCATG 132
Db 8 IleIleIleIleIleThrGlyGlySerGlyAlaGlyLysThrThrLeuThrGlnAsnIleLys 27
Qy 133 GAGTTGCTGGAGACAGAAAGAGTGGAAACAGCGGACGCGAAGTGTCTATCTGAGCCAG 192

Db 28 GluIlePheGlyGluAsp-----ValSerValIleCysGln 39
Qy 193 GACAGGTTCTACAG-----GTCTGACGCGACAGACAGAAAGCCCAAGGCTTGAAA 243
Db 40 AspAsnTyTyTyGluAspArgSerHisTyTyThrProGluGluArgAlaAsnLeuIle--- 58
Qy 244 GACAGTACATTTTACACATCCAGATGCTTGTATATATGATTTGATGACAGAGACTG 303
Db 59 -----TyrAspHisProSerAlaPheAspAsnLeuLeuIleSerAspIle 74
Qy 304 AAGACATCGTGGAGGCAAAACGATGAGGTGCCGACCTATGATTTGTGACACACTCA 363
Db 75 LysArgLeuLysAsnAsnGlnIleValGlnAlaProValPheAspPheValLeuGlyAsn 94
Qy 364 AGG---TTACAGAGACACACGCTGTCTACCTCGGACGATGTTCTTTGAGGCACTC 420
Db 95 ArgSerLysThrIleGluThrIleGluThrIleTyProSerLysValIleLeuValGluGlyIle 114
Qy 421 TTGGTGTCTTACACCGAGAGATCCGGACATGTTCCACCTGCGCTCTTCGTGAGACAC 480
Db 115 LeuValPheGluAsnGlnGluLeuAlaGlyAspLeuMetAspIleArgIlePheValAspThr 134
Qy 481 GACTCCGACGTCAGGCTGTCTCGAAGACTTCCGGACGTCGC---CGAGGAGGAGAC 537
Db 135 AspAlaAspGluArgIleLeuArgGlyMetValArgAspValGluGlnGlyAspSer 154
Qy 538 CTGAGACAGATTTGACGCGAGTACACACCTTGTGAAGCGCGCTTCGAGAGATTTGCG 597
Db 155 ValAspCysIleMetSerArgTyLeuSerMetValLysProMetHisGluLysPheIle 174
Qy 598 CTGCGCAAGAAAGATGCGCATGTGATCATCCACAGAGATGACAAATATGTTGCC 657
Db 175 GluProThrArgLysTyArgAlaAspIleIleValHisGlyAsnTyArgGlnAsnValVal 194
Qy 658 ATCAACCTGATCGTCGAG-----CACATCCAGAGACATTTGATGATGTCAG 702
Db 195 ThrAsnIleuSerGlnLysIleLysAsnHisLeuGluAsnAlaLeuGluIuSerAsp 213

RESULT 14
E72041
uridine kinase CP0011 [imported] - Chlamydia pneumoniae (strains CWL029 and AR39)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: E72041; E81622
R/Kalman, S.; Mitchell, W.; Marathe, R.; Lemmel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: E72041
A/Molecule type: DNA
A/Status: preliminary
A/Experimental source: strain CWL029
A/Cross-references: GB:AE001655; GB:AE001363; NID:g4377039; PIDN:AA018874.1; PID:g4377041
R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, J.
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: E81622
A/Molecule type: DNA
A/Status: preliminary
A/Experimental source: strain AR39, HL cells
A/Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AA037907.1; PID:g7188950
A/Genetics:
C/Genetics:
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 9.65e-25 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59

```

Query Match:      27.21%      Indels:      22
DB:               2          Gaps:         6

US-09-896-522-3 (1-834) x E72041 (1-222)

QY      73 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAATCATG 132
      8 11e1e61y1lethrglyserglyalaglylvsrthrthreuthrthnmaenllelye 27
DB      133 GAGTTGCTGGGACAGACGAGGTGAAACAGCGGACGGAGAGTGTCTGACCCAG 192
      28 G1u1lePhegllygluasp-----ValSerVal11leCysGln 39
QY      193 GACAGGCTTCACAG-----GTCTGACGGACAGACAGAGGACGCGCTTGAAA 243
      40 AspAsmtyrTyrrylsAspArgSerHisTyrrThrProGluGluArgAlaAsnLeulle-- 58
DB      244 GACAGATACATTTTGACATCCAGATGCTTGTATGATGATTTGATGACAGAGACTGTG 303
      59 -----TrpAspHisProAspAlaPheAspAsnAspLeuLeu1leSerAsp1le 74
QY      304 AAGAACATCTCGAGGCGCAAAACGGTGGAGGTGCCAATTATTTTGTGACACACTCA 363
      75 LysArgLeuLysAsnAsnGln1leValGlnAlaProValPheAspPheVal1leuGlyAsn 94
QY      364 AGG---TTACACAGACACACAGGTGTCTACCGTGGACGGAGTGTCTGTGAGCGCATC 420
      95 ArgSerLysThrGln1leGlnThr1leTyrrProSerLysVal1leuValGlnGly1le 114
DB      421 TTGGGTCTTACACGACGAGATCCGGGACATGTTCCACCTGCGCCTTCTGTGACACAC 480
      115 LeuValPheGluAsnGlnGln1leuValArgAspLeuMetAsp1leArg1lePheValAspThr 134
QY      481 GACTCCGACCTCGCGCTGTCTCGAAGATTCTCCGGGACGTGCG--CGAGGAGGAGAC 537
      135 AspAlaAspGlnArg1leLeuArgArgMetValArgAspValGlnGlnGlnGlyAspSer 154
DB      538 CTGGACACAGTTCTGACGCGGATCACACACCTGTGTAAGCGGCGCTTCGAGAGATTCTGC 597
      155 ValAspCysVal1leMetSerArgTyrrLeuSerMetValLysProMetHisGlnLysPhe1le 174
QY      598 CTCCCGACAAAGAGATGCGCGATGTCATGATCCACGAGAGTGGACATATGCTTGCC 657
      175 GlnProThrArgLysTyrrAlaAsp1le1leValHisGlyAsnTyrrArgGlnAsnValVal 194
DB      658 ATGACCTGATCTGTGAG-----CACATCCAGACATTTCTGAATGCTGAC 702
      195 ThrAsn1leuSerGlnLys1leLysAsnHisLeuGlnAsnAlaLeuGlnSerAsp 213

RESULT 15
G70101
uridine kinase (udk) homolog - Lyme disease spirochete
C/Spectes: Borrelia burgdorferi (Lyme disease spirochete)
C/Pate: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Nov-2000
C/Accession: G70101; T46979
R/Praeer: C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, S.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A/Authors: Smith, H.O.; Venter, J.C.
A/Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A/Reference number: A70100; PMID:98065943; PMID:9403685
A/Accession: G70101
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-207 <KLE>
A/Cross-references: GB:AE001115; GB:AE000783; NID:g2687879; PIDN:AA66392.1; PID:g268788
A/Experimental source: strain B31
R/Bouraux-Ende, C.; Margarita, D.; Gilles, A.M.; Barzu, O.; Gitons, I.S.
FEMS Microbiol. Lett. 151, 257-261, 1997
A/Title: Borrelia burgdorferi uridine kinase: an enzyme of the pyrimidine salvage pathwa
A/Reference number: Z24339; PMID:97372541; PMID:9228761
A/Accession: T46979

```

```

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 2-207 <BOU>
A/Cross-references: EMBL:X97449; NID:g1405440; PIDN:CAA66081.1; PID:g1405442
A/Experimental source: strain HB19, ssp. sensu stricto
C/Genetics:
A/Gene: udk
C/Superfamily: uridine kinase

Alignment Scores:
Pred. No.:      1,26e-24      Length:      207
Score:          410.50      Matches:     79
Percent Similarity: 59.81%      Conservative: 49
Best Local Similarity: 36.92%      Mismatches: 59
Query Match:    27.11%      Indels:     27
DB:             2          Gaps:         4

US-09-896-522-3 (1-834) x G70101 (1-207)

QY      73 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAGATGCACTGTGTGAGAAATCATG 132
      5 11e1e61y1lethrglyserglyalaglylvsrthrthreuthrthnmaenllelye 24
DB      133 GAGTTGCTGGGACAGACGAGGTGAAACAGCGGACGGAGAGTGTCTGACCCAG 192
      25 GlnPhe1leProGlu-----PheVal1leu1leSerGln 35
QY      193 GACAGGTTTACAGGTTCTGACGCGGACGAGACGAGGCCAAGGCTTGAAAGACACTAC 252
      36 AspAsnTyrrTyrrLysSerVal-----GlyAspTyrr 45
DB      253 -----AATTGACATCTCCAGATGCTTGTATGATGATTTGATG 291
      46 GlnHisGlnPheSerLysValAsnPheAspHisProAspAlaPheAspAsnAspLeuPhe 65
QY      292 CACAGGACTCTGAGAACATCTGTGAGGCGCAAAACGGTGGAGGTGCCAATTATGATTTT 351
      66 TyrrGlnHisLeuLysAsnLeuLysValAsnSerPro1leAspMetProLeuTyrrAspPhe 85
DB      352 GTGACACACTCAAGTTTACAGAGACACACGAGTGTCTACCTGCGGACGTGTCTGTTT 411
      86 1leAsnHisLysArgGlnLeuLysThrVal1leuValAlaProThrProVal1leVal 105
QY      412 GAGGAGATCTTGCTGTCTACAGCCGACGAGATCCGGGACATGTTCCACCTGCGCCTTTC 471
      106 GlnGly1leMet1lePheValGlnGluArgValArgAsnLeu1leAspLeuLys1leTyrr 125
DB      472 GTGGACACGACACTCCGACGTCGAGTGTCTCGAAGAGTCTCCGGGACGTG--CGCCGA 528
      126 1leAspThrProAsnAsp1leArgPhe1leArgValGlnArgAsp1leSerLysArg 145
QY      529 GGGAGGACCTGGAGACGATCTGACGACGATCACACACCTTGTGAAGCGGCGCTTGAG 588
      146 GlnArgThrValGlnSerVal1leAspGlnTyrrLeuAsnThrThrArgTyrrGlyTyrr 165
DB      589 GAGTTGCTGCTCCGACAAAGATGCGCGATGATATATCCACGAGAGTGGACAT 648
      166 ArgPhe1leGlnProThrLysGlnTyrrAlaAsp1le1leProGlnGlnGlnHisAsn 185
QY      649 ATGGTGGCATCAACTGATCTGTGACGACGATCATTCACAGAGACATT 690
      186 AspLysAlaLeuTyrrValLeuSerThrPheLeuLysSerLeu 199

```

Search completed: November 25, 2003, 07:40:43
 Job time : 26.008 secs

This Page Blank (uspto)

	CURRENT FILING DATE:	2001-08-10
	PRIOR APPLICATION NUMBER:	PCT/US00/05988
	PRIOR FILING DATE:	2000-03-08
	PRIOR APPLICATION NUMBER:	60/124,270
	PRIOR FILING DATE:	1999-03-12
	NUMBER OF SEQ ID NOS:	1890
	SOFTWARE:	Patentcin Ver. 2.0
	SEQ ID NO	1160
	LENGTH:	337
	TYPE:	PRT
	ORGANISM:	Homo sapiens
	FEATURE:	
	NAME/KEY:	SITE
	LOCATION:	(38)
	OTHER INFORMATION:	Xaa equals any of the naturally occurring L-amino acids
	NAME/KEY:	SITE
	LOCATION:	(46)
	OTHER INFORMATION:	Xaa equals any of the naturally occurring L-amino acids
	NAME/KEY:	SITE
	LOCATION:	(155)
	OTHER INFORMATION:	Xaa equals any of the naturally occurring L-amino acids
	NAME/KEY:	SITE
	LOCATION:	(169)
	OTHER INFORMATION:	Xaa equals any of the naturally occurring L-amino acids
	US-09-925-300-1160	
	Alignment Scores:	
	Pred. No.:	9.18e-68
	Score:	897.50
	Percent Similarity:	73.31%
	Best Local Similarity:	64.53%
	Query Match:	59.28%
	DB:	10
	Gaps:	6
US-09-896-522-3 (1-834) x US-09-925-300-1160 (1-337)		
QY	3 GGGTTGGGGGAGCGA-----AGACTGGAGAAGCCCGCGCGAGGCCCA---	50
Db	57 GLVALAGLYGLUALRGLASERVALARGHRLYSERGLYARGARGLYALAAASN	76
QY	51 -----CCGCACCACCA-----GGGCC	68
Db	77 HIEGLVIRGLYLNRGLAASPROMLAGLUPROPOLALAIGLNATGATGATGALA	96
QY	69 CTTCCTCATAGGGGTGAAGCGGCGCAGTCCAGCGGGAAGTCGACCGTGCTGAGAAAT	128
Db	97 LEPRDLYARARGHIS--GLGYLTRHLASERGLYLSERSESERVALCYSLALAYSIL	116
QY	129 CATGAGATTGCTGGGACAKAGACAGAGTGAACAAGCGGACGGGAAGGTGTCCTCTGAG	188
Db	116 EValGINLEULeuGLYGlnAsnGLUValAspTYrFArgINLysGLInLysGLInValIleLeuSe	136
QY	189 CCAAGACAGGTTCTCAAGAGTCTGACGGGACAGACAGAGGCCAACGGCTTGAAAAGACA	248
Db	136 rGlnAspSerPheTYrFArgVALLeuThrSerGLNGlnMysAlaYsaIleuLys**GI	156
QY	249 GTACATTTTGAACAATCCAGATGCTTTGATATGATTTGATGACACAGACTCGAAGA	308
Db	156 nPrElaAnPheArHisPROMerPalAPhaAspAsnGLU**IleuLysThrLeuLysGI	176
QY	309 CATCGTAGGSCAAACCGGTGAAGTCCGACCTATGATTTTGTACACACTCAAGGTT	368
Db	176 uIlEtTrgInGLyLysThrValGlnLleProValTYrFAspPheValSerHisSerArgLy	196
QY	369 ACCAGAGACAGAGGTGCTCACCGCGGACAGTGGTCTGTGGAGGCAATTTGGGTT	428
Db	196 sGIuGLInThrValTYrPROlAAspValValLeuPheGLuLYlleuAlaPh	216
QY	429 CTACAGCCAGAGATCCGGACATGTTCACTTCAGCGCTTCGTGGACACCACTCCGA	488
Db	216 eTyrsErGlnGLuValArgAspLeuPheGlnMeTLysLeuPheValAspThrAspAlas	236
QY	489 CGTCAGGCTGTCTGAAAGATTTCTCCGGACGTG---CGCCAGAGGAGGACCTTGAGCA	545


```
Db      |||...|||
236 ptnrArgLeuSerArgArgValLeuArgAspIleSerGluArgGlyArgAspLeuGlu 256
Qy      546 GATTTCAGCCAGTACACCACTTCGTGAAGCCGCGCTTCGAGGAGTTTCGCTGCCGAC 605
Db      256 nileleSerGlnTyrIleThrPheValysProAlaPheGluGlnIlePheCysLeuProth 276
Qy      606 AAGAAGTATGCGCATGATGATCCAGAGGAGTGAGACATATGATGTCCTCCATCAACT 665
Db      276 rlyblyrAlaAspAlaIleleProArgGlyAlaAspAsnLeuValAlaIleAsnle 296
Qy      666 GATCGTCAGACATCCAGACATTCGAAATGTGATCTGCAATATGACACCGAGAGG 725
Db      296 uileValGlnHietIleGlnAspIleLeuAsnGly----- 307
Qy      726 GTCCATAGGCGGAGCTACAAAGCGGACCTTTCTGAGCCAGGAGCAACCTCGGATGCT 785
Db      308 -----GlyProSerIysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrTh 322
Qy      786 GACCTCGGAAACGGTCACATTTGAGTCCAGAGAGACCCAC 831
Db      322 rProSerArgIysArgGlnAlaSerGluSerSerArgProHis 337

RESULT 3
US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Alignment Scores:
Pred. No.: 5,34e-28 Length: 125
Score: 428.50 Matches: 84
Percent Similarity: 76.80% Conservative: 12
Best Local Similarity: 67.20% Mismatches: 28
Query Match: 28.30% Indels: 1
DB: 9 Gaps: 1

US-09-896-522-3 (1-834) x US-09-896-522-5 (1-125)
Qy      460 CTGCGCTCTTCGTGGACACCGACTCCGAGTCAAGGCTGTCTGGAAGTTCCTCCGGAC 519
Db      1 LeuLysIlePheValaPthrAspAlaAspValaArgLeuIleArgGlyIleLysArgAsp 20
Qy      520 GTG---CGCGAGGAGGAGCGTGAAGACATTCGACGAGTACACCACTTCGTAAG 576
Db      21 ValaAsnGluArgGlyArgAspIleGlnSerValIleGlnGlnIleTyrMetLysPheValLys 40
Qy      577 CGCGCTTCGAGAGTTCCTGCTGCGACAAAGATGCGGATGATGATCATCCCA 636
Db      41 PrcMetTyrGlnGlnPheIleGlnProThrIleTyrAlaAspIleIleIleProArg 60
Qy      637 GGAAGTGAACAATATGCTTGCATCAACCTATGTCGAGACATCCGACATTCGAAT 696
Db      61 GlyGlyAspAsnHisValaAlaIleAspLeuIleValGlnHisIleGlnSerIleLeuAsn 80
Qy      697 GGAGACATTCGAATATGACCGAGAGGCTCAATGCGGAGACTACAGCGAAGCTTT 756
```

```
Db      |||...|||
81 GluGlyLeuSerSerGlnHisThrAsnTyrMetValaAsnArgSerTyrIysArgThrPhe 100
Qy      757 TCTGAGCCAGGAGGAGCAACCTCGGAGTGTGACTCTGCGAAACGGTCACATTTGGAGTCC 816
Db      101 SerGluProGlyAspHisProGlyTyrThrProSerGlyLysArgGlnHisLeuGluSer 120
Qy      817 AGCAGCAGACCCAC 831
Db      121 SerSerArgProHis 125

RESULT 4
US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Alignment Scores:
Pred. No.: 1.86e-19 Length: 60
Score: 327.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.60% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-3 (1-834) x US-10-029-386-28765 (1-60)
Qy      652 GTTGCCATCAACCTGATCGTCGACGACATCCAGACATTCGAATGTCGACATTCGCAAA 711
Db      1 ValaAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 20
Qy      712 TGGACCGAGAGGAGTCCATATGAGCGGAGCTTCAACCGGACCTTTTCTGAGCCAGGAGAC 771
Db      21 TrpHisArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 40
Qy      772 CACCCCTGGATGTCGACCTTCGGCAACCGTCACATTTGATGATCCAGAGACACCCAC 831
Db      41 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 60

RESULT 5
US-09-896-522-6
; Sequence 6, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE OF INVENTION: USES THEREOF
```

```
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-6

Alignment Scores:
Pred. No.:      8,17e-19      Length:      124
Score:          320.50        Matches:      62
Percent Similarity: 67.19%    Conservative: 24
Best Local Similarity: 48.44%  Mismatches:    37
Query Match:    21.17%       Indels:       5
DB:             9            Gaps:         2

US-09-896-522-3 (1-834) x US-09-896-522-6 (1-124)

QY 73 CTGATAGGGGTGAGCGGGGACCTGCCAGCGGAAAGTCGACCGTGTGTGAGAAATCATG 132
DB 1 I l l e e g l y l e a l a g l y s e r g l y s e r t h r l l e a l a g l y s i l e v a l 20
QY 133 GAGTGTCTGGACAGAACGAGGTGAAACAGCGGACGCGGAGGAGTGTGATCTCTAGCCAG 192
DB 21 G l u t e t l e u a n ----- l y e p r o g l y n g l u l y s v a l l e i l l e s e r c l n 36
QY 193 GACAGGTTCTTAAAGTCTTGACGCGACAGGAAAGCGCAAGCTTGAAGAGACAGTAC 252
DB 37 A s p a n t y r t y l y s a s p l e u s e r c l u d e u a s p m e t g l u d a r g l y s l u a s n a n t y x 56
QY 253 AATTTGACCATCCAGATGACCTTGTGATTAATGATTTGAGCAGACAGACTCTGAAAGATC 312
DB 57 A s n h e a s p h i s p r o a s p a l a p h e a s p h e a s p l e u e u t y g l u n h i s l e u l y s **L e u 76
QY 313 GTGAGGCGCAAAAGGTGAGAGTCCGACCTATGATTTTGTGACA---CACTCAAGGTTA 369
DB 77 L y a s n e n g l y s s e r v a l g l u v a l p r o l l e t y r a s p h e u y s t h r h i e h i s A r g d i r l y s 96
QY 370 CCAAGACCAACGCGGTCTTACCTGCGGACGCTGCTTGTGAGGCGATCTTGATTC 429
DB 97 A s p g l u t h r v a l t h r l l e g l u p r o l a l a s p a l l e i l l e u g l u g l y l e t y r a l a l e u 116
QY 430 TACAGCCAGGATCCGGGACATG 453
DB 117 T y r a s p g l u a r g l i e a r g a s p l e u 124

RESULT 6
US-09-896-522-4
; Sequence 4, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-4 (1-834) x US-09-896-522-4 (1-231)

QY 73 CTGATAGGGGTGAGCGGGGACCTGCCAGCGGAAAGTCGACCGTGTGTGAGAAATCATG 132
DB 1 I l l e e g l y l e a l a g l y s e r g l y s e r t h r l l e a l a g l y s i l e v a l 20
QY 133 GAGTGTCTGGACAGAACGAGGTGAAACAGCGGACGCGGAGGAGTGTGATCTCTAGCCAG 168
DB 21 S e r l l e p h e d l y a r g l u g l y a l p r o l a l a g l y l l e g l u d l y a n p r o a s p e r a n 40
QY 169 CGAAGGTGTCTATCTGACCCAGACAGGTTTACAGGCTCTGACGCGACAGCAAG 228
DB 41 T h r g l y a s p s e r p h e l e u a r g l e u a s p a r g p h e t y m e c a s p l e u n h i s l e u g l u a s p a r g 60
QY 229 GCCAAGGCTTGAAGAGACGATCAATTTTGACCATCCAGATCCCTTGATTAATGATTTG 288
DB 61 L y a r g l a l a l y a n l y s h i s t y s e r p h e s e r p r o g l u l a s n a s p h e a s p l e u 80
QY 289 ATCCACAGCACTCTGAAGACATCGTGAAGGCGCAAAAGGTGAGGCTGCGCACTATGAT 348
DB 81 L e u t y r g l u v a l a h e l y s g l u e u l y s g l u g l y l y s e r V a l a s p r o l l e t y r a n 100
QY 349 TTTGTACACACTCAAGGTTACCAAG-----ACCAAGGTGCTTACCT 393
DB 101 H i e v a l t h r c l y l u a r g a s p r o a s p a r g l y n g l u p r o g l y t h r p h e t h r a s p t r p r o 120
QY 394 -----CGGACGCGGTTCTGTTGAGGACCTTGTGATCTGATCTTACCCAG 438
DB 121 G l u l e u l l e g l u g l y l a s p a l l e u a l l e g l u l y n e u n i s a l a l e u t y r a s p g l u 140
QY 439 -----GAGATCCGGGACATGTTCCACTGCGCTTCTCTGACACCGACTCGAC 489
DB 141 A r g l u v a l a s n a l a g l u n l e u a s p l e u l y l l e t y r a l a s p r o a s p i l e a s p 160
QY 490 GTCAAGCTGTCTGAAGAGTTCTCCGGACGTG---CGCCGAGGAGGACCTTGAGCAG 546
DB 161 L e u g l u e u a l a r g y s i l e g l n a r g a s p m e t a l a g l u a r g l y h i s s e r l e u g l u g l y 180
QY 547 ATTCTGACGAGTACACACCTTCGTGAAGCGCGCTTGAGAGGAGTTCGCTCCGCGACA 606
DB 181 V a l l e u a s p s e r l l e g l u l y s a r g a r g l y s p r a s p t y r v a l a n t y r l l e a l a p r o g l n 200
QY 607 AAGAAGTATGCCGATGATCATCCACGA 636
DB 201 P h e s e r t y r l a a s p l e u l l e i l l e g l n a r g 210

RESULT 7
US-09-975-719-369
; Sequence 369, Application US/09975719
; Publication No. US20030022349A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Rahme, Laurence G.
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361003
; CURRENT APPLICATION NUMBER: US/09/975,719
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 09/199,637
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: US 60/066,517
; PRIOR FILING DATE: 1997-11-25
```

```
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-975-719-369

Alignment Scores:
Pred. No.: 4.55e-05      Length: 265
Score: 160.00           Matches: 87
Percent Similarity: 35.62% Conservative: 22
Best Local Similarity: 28.43% Mismatches: 107
Query Match: 10.34%      Indels: 90
DB: 11                  Gaps: 17

US-09-896-522-3 (1-834) x US-09-975-719-369 (1-265)
QY      820 TGGTGAAGTCCAAATGTGACCGTTTGGCAGAGGTACAGATCCGAGGGTGGTCCCTGGCT 761
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      11  CysArgSerProAlaValProArgCysCysArgSerIleYs--GlyArgAlaAla 29
QY      760 CAGAAAAGTCCGCT-----TGTAGCTCCGCCAT-----731
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      30  GluSerArgSerAlaProTyrProAlaProCysAlaIleAlaIleAlaIleProArgThrCys 49
QY      730 -----TGACCCCTCCGCTGCGCATTTGCAGATGCACATTCAGAAATGCTCGATGT 677
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      50  ArgGlySerSerAspLeuSerProValProArgAlaValProGlyAlaAlaSerThr 69
QY      676 GGTGACAGATCAGATGATGAGCAACATATTTGTCATCTCTCGGGAGTATCATCATCGG 617
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      70  ThrArgArgSerArgAlaGluProProAlaAlaProCys-----Arg 83
QY      616 CATATCTTCTTTGTCGCGCAGCAAGAACTCTCGAAGCCGCGCTTCACGAAAGTGGTACT 557
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      84  HisGlySerArgHisAla--ArgSerProArgArgProLeuSerArgArg-----Arg 100
QY      556 GCGTCAGAAATCTGCTCCAGATGCTCCCTCGCGGCGACAGTCCCGAAGAACTTTTCAGACA 497
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      101  AlaThrSerArgAlaPro--ProSerValHisAla-----111
QY      496 GCGTGAAGTGGAGTGGTGCACGAAAGGCGCAGTGAACATGTCCCGATCTCT 437
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      112  -----ThrArgAspTyrArgGlyLeuAlaIleAlaIleProAlaMetSerAlaPro 127
QY      436 GGGCTGTAACAACCAAGATCCCTCAACAAGAACACAGTCCGAGGAGTACCAACCGTGG 377
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      128  Gly-----GluSerCysValAlaHisTyr--ProArgProThrCysArgProPro-- 142
QY      376 TCTCTGGTAACCTTGAAGTGTGTCAAAATCATATGAGTGGCACTCCACCGTTTGCCCT 317
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      143  -----AlaGlySerLeuArgProAlaArgAlaProProAlaArgPro 156
QY      316 CCAAGATGTTCTTCAGAGTCTCTGATCAATATCATTAACAAGGCAATCGATGTGCA 257
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      157  ProArgSerAlaLeuProSerThrAla-----GlyCys 167
QY      256 AATTGTAAGTCTCTTCAGAGGCTTGGCTTGTGCTGCGG-----TCAGA 209
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      168  GlyCysVal-----ProThrAlaValProArgSerGlyArgProGly 181
QY      208 CCTTGTGAAC-----197
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      182  ProAlaSerThrGlyAlaArgGlyArgProArgAspArgAlaArgSerArgArgAlaAsn 201
QY      196  ---TGTCCTGCTCAGAGATCAACCTTCGCGTGGCTGTCACCTGCTGTCGCA 140
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      202  AlaAlaProIleProGlyArgSerProAlaThrSerAla--ProGlyArgSerProPro 220
QY      139  GCAACTCCATGATCTTCTCAACAACGCTGCATCTCCGCTGGCAGTGGCCGCGCTCAACC 80
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      221  GlyCys-----SerAlaThrAspArgAlaAlaIleArgArgSerArg 233
```

```
QY      79  CTATCAGAAAGGCGCGCTGTGTGCGACGATCGGCGCTCGGCGCGGCGCTCTCGACACTTT 20
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      234  LeuProProArgThrProHisAlaHisSerArgProAlaProProIleSerProAlaVal 253
QY      19  CGCCTCCCGCCGAAGCCA 2
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      254  ArgArgAlaProAlaPro 259

RESULT 8
US-09-738-626-4609
; Sequence 4609, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4609
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4609

Alignment Scores:
Pred. No.: 0.00014      Length: 312
Score: 154.50           Matches: 61
Percent Similarity: 44.64% Conservative: 39
Best Local Similarity: 27.23% Mismatches: 87
Query Match: 10.20%      Indels: 37
DB: 10                  Gaps: 12

US-09-896-522-3 (1-834) x US-09-738-626-4609 (1-312)
QY      67  CCTTCTGATAGGGGTGAGCGCGGCACTGCGCAGCGGAAGTCAACCGTGTGTGAAG 126
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      90  ProPheValIleGlyValAlaGlySerValAlaValGlySerThrThrAlaGluLeu 109
QY      127 ATCATGGAAGTGTGCTGGGACAGAAAGAGTGAACACGG-----CAGCGGAAGGTG 177
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      110  LeuGlnValLeuLeu-----GlnArgTyrAsnSerHisProArgVal 123
QY      178 GTGATCCTGAGCAGGACAGATTTCTACAGGTCTTACGGCAGAGCAAGCAAGCGCAAGGCC 237
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      124  AspLeuValThrThrAspIlePheLeu--TyrProGlyAlaGluLeuIleArgArgGly 142
QY      238 TTGAAAGACAGTACATTTTGAACATCCAGATGCTTGTGATTAATGATTTTGAACAGAG 297
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      143  LeuMetSerArgIleGlySerHe-----ProGluSerTyrAspGlnArgAlaLeuLeuArg 160
QY      298 ACTCTGAAGAACATCTCGAGGCGCAAA--ACGGTGAAGGTGCCGACCTATGATTTTGTG 354
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      161  PheValThrAspValIleGlySerGlyLeuGluValAsnAlaProValIleSerHisThr 180
QY      355 ACACACTCAAGGTACCA--GAGACCAAGGTGTACTCCGCGGACGTGTCTGTT 411
```



```

1  APPLICANT:  IKEGA HARUO
2  APPLICANT:  ISHIKAWA, JUN
3  APPLICANT:  HORIKAWA, HIROSHI
4  APPLICANT:  SHIBA, TADAYOSHI
5  APPLICANT:  SAKAKI, YOSHIYUKI
6  APPLICANT:  HATTORI, MASAHIRA
7  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
8  FILE REFERENCE: 249-262
9  CURRENT APPLICATION NUMBER: US/10/156,761
10 CURRENT FILING DATE: 2002-05-29
11 PRIOR APPLICATION NUMBER: JP 2001-204089
12 PRIOR FILING DATE: 2001-05-30
13 PRIOR APPLICATION NUMBER: JP 2001-272697
14 PRIOR FILING DATE: 2001-08-02
15 NUMBER OF SEQ ID NOS: 15109
16 SEQ ID NO 12495
17 LENGTH: 329
18 TYPE: PRT
19 ORGANISM: Streptomyces avermitilis
20 US-10-156-761-12495

```

```
Alignment Scores:
```

Pred. No.:	0.000417	Length:	329
Score:	149.00	Matches:	62
Percent Similarity:	43.67%	Conservative:	38
Best Local Similarity:	27.07%	Mismatches:	89
Query Match:	9.84%	Indels:	40
DB:	15	Gaps:	11

US-09-896-522-3 (1-834) x US-10-156-761-12495 (1-329)

US-10-078-547-24

Alignment Scores:

Pred. No.: 0.0132 Length: 507
 Score: 132.00 Matches: 79
 Percent Similarity: 31.29% Conservative: 13
 Best Local Similarity: 26.87% Mismatches: 124
 Query Match: 8.53% Indels: 78
 DB: 14 Gaps: 12

US-09-896-522-3 (1-834) x US-10-078-547-24 (1-507)

```

QY 801 CCGTTTCCAGAGTGCATCCAGGGTGTGCTCCCTGCTCAGAAAGGCGCTTGTGTA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 PrLIeInserSerLeuHISaNaArgGlySerProProValProGlyGlyProAlaGln 215
QY 741 GCTCCGCCATTGGACCCCTCCGCTGCTCCATTGGCAGATGTCACCATTCAGAAATGTCCTG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 ProSerPro---GlyProThrProProPhe----- 225
QY 681 GATGTGCTGCACGATCAGTTGATGGACACCATATTGTCATCTCTGCTGGATGATCAC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 -----ProGlyAsnaArg---GlyThrAlaLeuGlyGlyGlySeri 238
QY 621 ATCGGCATATCTTTGTGCG-----GCAGGCAAGAACTCCTCGAAGGCGCG 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 leArgInserProLeuSerSerSerSerProPheSerAsnaArgProProLeuProProT 258
QY 576 CTTACGAGAGTGTGTACTGCGTCAGAAATCTGCTCCAGGTCCTCCCTCGGCGCAGCTC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 hrProSerArgAlaLeuAspAspIysProProProProProProProValGlyAsnaArgP 278
QY 516 CCGGAGAACTCTTGCAGACAGCTGAGTCCGCTGCTCCAGAAAGGCGCAGCTG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 ro-----SeriLeHISaArgGlyAlaValProProProProProGlnA 292
QY 456 GAACATGTCCTGGATCTCTGCTGTGAGAACCAAGATGCGCTCAAAACAGAACCAAGCTC 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 snAsnIysProProValProSer-----ThrProArg---ProSerAlaProHISaArgP 309
QY 396 CGCAGAGTAGAACCCGCTGTCTCTGTGTAACTTTGAGTGTGTCAAAATCATAGTACGG 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 roHISaLeuArgProPro-----ProProSerArgProG 320
QY 336 CACCTCCACCGCTTTTGCTCCAGATGTTCTTCAAGTCTCTGTCATCAAAATCATATTC 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 IyProProProLeu---ProProSerSerSerGlyAsnaArgGlyThrProArgLeuProG 339
QY 276 AAAGCATCTGATGTGTCAAATTTGACT-----G 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 InArGAsnLeuSerLeuSerSerSerThrProProLeuProSerProGlyArgSerGlyP 359
QY 246 TCCTTTCAAGGCTTTGGCTTGTGCTGCTGCGTCAGAGACCTTGTAGAACCTGTCTGGCT 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 roLeuProProProValProSerGlnArgProProProProValArgAspProProGlyA 379
QY 186 CAGAGTACCACTTCGCTCGCTCGCTGTTCACCT----- 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 rgSerGlyPro-----LeuProProProProProValSerArgAsnGlyS 394
QY 151 -----CGTTGTGTCAGAACTCCATGATCTTCTACACAGGTCGACTTCCGCT 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 erThrSerArgAlaLeuProHISaThrPro-----GlnL 405
QY 99 GGCAGTCCGCGCTGCACCCCTATCAGAAAGGCGCTGTGCGGACGATCGGCTCGG 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 405 euProSerArgSerGlyValAspSerProArgSerGlyProArgProProLeuProProA 425
QY 39 CGGCGGCTCTCGCAGCTTCGCTCGCTCCGCGCAAGCA 2
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 425 spArgProSerAlaGlyAlaProProProProPro 437
  
```

RESULT 14

US-10-081-872-208

; Sequence 208, Application US/10081872
 ; Publication No. US20030125534A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Callen, Walter
; APPLICANT: Richardson, Toby
; APPLICANT: Frey, Gerhard
; APPLICANT: Short, Jay M.
; APPLICANT: Mathur, Eric J.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Kerveno, Janne S.
; APPLICANT: Slupska, Malgorzata
; TITLE OF INVENTION: ENZYMES HAVING ALPHA AMYLASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: 09010-108001
; CURRENT APPLICATION NUMBER: US/10/081,872
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,495
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/270,496
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/291,122
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Environmental
  
```

US-10-081-872-208

Alignment Scores:

Pred. No.: 0.017 Length: 439
 Score: 130.50 Matches: 69
 Percent Similarity: 35.38% Conservative: 23
 Best Local Similarity: 26.54% Mismatches: 99
 Query Match: 8.44% Indels: 69
 DB: 15 Gaps: 16

US-09-896-522-3 (1-834) x US-10-081-872-208 (1-439)

```

QY 709 TGCAGATGTACCATTCAGAAATGCTGATGTGTGCACGATCAGTTGATGAGCAACA 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CysSerCysGlnArgProArgProSerCys---SerArgAlaGlySerArgArgPro 200
QY 649 TATGTTCACCTCTCGTGGAGATGATCAATCGGCATATCTTTGTGCGCAGGCAAGCT 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 Trp-----ArgArgSerSerArgProSerGly----- 209
QY 589 CCTCGAAGCGCGCTTCAGAAAGTGTGTACTGCTCAGAAATCTGCTCA----- 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 -----ValHISaGlnArgTrpCysProSerThrArgGlnArgProSerArgPro 225
QY 538 -----GATGCC-----TCCCTCGGCGCAGCTGCC 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 ThrSerAlaSerProArgProThrLeuArgGlyProSerArgSerGlnSerAlaArgHIS 245
QY 514 GAGAACTTTGAGACAGCTGACGTGAGT-----CGGTGTCAGAAAGGCGCA 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlnArgArgCysSerLeuGlyArgArgArgSerSerHISaArgSerProArgAlaSerAla 265
QY 460 GGT-----GGAACATGTCCTCGGATCTCCCTGCGTGTGAACACCAACATGC 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 GlyProSerSerSerArgGlyLeuCysLeuGlySer-----LeuGlnMetCys 281
QY 415 CCTCAA---ACAGAACCAAGCTCGCAGAGT---AGACCAACGATGCTCTGTGAACCTTG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 ProArgHISerThrProArgTrpGlyGlySerArgGlySerTrpGlnTrpIleCys--- 300
QY 361 AGTGTGTCACAAAATCATAGTGTGCGACCTTCACCGTTTGGCTTCACGATGTTTCTTA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 -----ProArgProProLeuArgSerProSerArgCysSerPro 313
  
```

```

QY 301 GAGTCCTGTGATCAATATCATTAACAGCATCTGATGTCATAAATTGACTGCTT 242
Db 314 GlnArgThrIleGlySerThrArgGlyLeuArgLeuArgGlyLeuArgCys----- 330
QY 241 TCAGAGCCTTGACCTTCTGCTCTGCGCTGACGACCTTTAGAACCTGTCTGCGCTCAGGA 182
Db 331 -----ProLeuArgProLeuCysArgArgHisGlyProCysLeuSerCysSerArgAlaPro 348
QY 181 TGACCACTTCCGCGCTGTTCCACCTGCTGTCTGTCACCAAGCATCCATGATCTTCT 122
Db 349 AlaTrpSerGlnSerAlaSerLeuPro-----PhePro-----SerGly 361
QY 121 CACACACGCTGACTTCCGCTGCGATGCGCGCTCACCCCTATCAGAGAGGCGCGCT 62
Db 362 ArgThrHisArgGlyGlnArgSerArgArgGlyArgSerProSerAsnArgArgPro 381
QY 61 GGTGCGACGCTCGCGCTCCGCGCGGCGCTCTGCGACTTTCGCTC--CCGCCGAG 5
Db 382 -----CysProCysSerProGlyGlnSerIleTrpArgIlePheProArg 397

```

RESULT 15

```

US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-10-024-368-5

```

```

Alignment Scores:
Pred. No.: 0.0305 Length: 1207
Score: 129.00 Matches: 67
Percent Similarity: 33.33% Conservative: 14
Best Local Similarity: 27.57% Mismatches: 72
Query Match: 8.34% Indels: 90
DB: 15 Gaps: 14

```

US-09-896-522-3 (1-834) x US-10-024-368-5 (1-1207)

```

QY 583 AGGCCGCTTACGACAGGTGTGTACTGCGTCAGATCT----- 545
Db 241 ArgProAla--LysGluProLeuAlaIleSerGlnMetLeuGlyLeuAlaPro 259
QY 544 -----GCTCAGAGTCCCTCCCTCGGCGACGTCGCCGA 512
Db 260 GlyGlyHisLeuGlnGlnAlaCysAspAlaGluGlyProSerLeuHisGlnArgAspGly 279
QY 511 GAACCTTCGAGACAGCTGACGTGCGAGTGTGCCACGAGAGCGCAGGT----- 458
Db 280 Glu-----ThrGlyAlaGlyArgGln 286
QY 457 ---GGAACATGTCCCGATCTCTGCTGTAGAACACCAAGATGCCCTCAACAGAACCA 401
Db 287 GlnAsnLeuCysProValPheLeuGly-----TyrProAspThrValPro 301
QY 400 CGTCCG-----CAGGATAGACACCGGTGTCTCTGGAACCTTG----- 362
Db 302 ArgThrProTrpProSerCysProProGlyLeuValHisThrLeuGlyAsnValTrpAla 321
QY 361 -----AGTGTGTCACAAAATCATATAGTGGACACTCCA-----CCGTTTTCGCT 317

```

```

Db 322 GlyProGlySerAsnSerPheGlyTyrGlnLeuGlyProProValThrProArgCysPro 341
QY 316 -----CCAGAGATTTCTTACAGACTCTGTGCATCAAAATCATTAACAGCATCTG 266
Db 342 SerProGlyProProThrProProGlyGlyCysCysSerSerHis----- 356
QY 265 GATGTCAAAATTGTACTGTCTTTCAGAGCCTTGACCTTCTGCTGCGCTCA----- 212
Db 357 -----LeuProAlaArgGln 361
QY 211 -----GACCTTGTAACCTGTCTGCTGATGATGACCACTTCGCTGCGCT 161
Db 362 GlyAspProGlyProCysArgGlyCysGlnAspSer-----ProGlnGlySerSer 379
QY 160 GTTCCACTGCTGTCTGTCCA-----GCACTCCATGATCT 125
Db 380 GlyProGlyGlnSerSerGlnArgAsnLysAlaGlySerArgAlaSerProSer 399
QY 124 TTCACACACAGTTCGACTTCCGCTGAGCGCGCTCACCCCTATCAGAGAGGCC 65
Db 400 HisHisThrIleLeuIleLysIleLysThrTrpLeuThrArgHisSerGlnInPheGluCysPro 419
QY 64 GCTGTGTCGAGACGCTCGG-----CCTCCGCGCGCGGCTTTCGACGCTTCGCTCCCG 11
Db 420 GlyGlyCysProGlyLysGlyGlnSerProAlaThrGlyLeuArgAlaLeuLysArgAla 439
QY 10 CCGAAGCCA 2
Db 440 GlySerPro 442

```

Search completed: November 25, 2003, 07:50:36
Job time : 98.057 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:27:32 ; Search time 10.688 Seconds
(without alignments)
6603.180 Million cell updates/sec

Title: US-09-896-522-3

Perfect score: 1514

Sequence: 1 atggctctgcgcggagcgca.....ccagcagcagaccactga 834

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/rnat_2112003.184106.2885/app_query.fasta.1.2830
-DB=Issued Patents AA -QFMT=fasta -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CEN_1.1.29.@rnat_2112003.184106.2885 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1363	90.0	260	4	US-09-536-647-2
2	1325	87.5	277	4	US-09-536-647-3
3	451	29.8	216	4	US-09-134-001C-3618
4	434	28.7	210	4	US-09-107-532A-5024
5	412	27.2	222	4	US-09-198-452A-783
6	164	10.6	1150	4	US-09-252-991A-24671
7	160	10.3	365	4	US-09-199-637A-369
8	158.5	10.2	274	4	US-09-252-991A-25569
9	157.5	10.2	247	4	US-09-252-991A-26899
10	157	10.4	664	4	US-09-252-991A-31116
11	156	10.1	335	4	US-09-252-991A-24899
12	156	10.1	395	4	US-09-252-991A-22433

c 13	155.5	10.1	265	4	US-09-252-991A-21436	Sequence 21436, A
c 14	153	9.9	663	4	US-09-252-991A-30843	Sequence 30843, A
c 15	152.5	10.1	582	4	US-09-252-991A-24055	Sequence 24055, A
c 16	151.5	9.8	472	4	US-09-252-991A-31978	Sequence 31978, A
c 17	150.5	9.7	272	4	US-09-252-991A-31371	Sequence 31371, A
c 18	150	9.7	266	4	US-09-252-991A-19128	Sequence 19128, A
c 19	149.5	9.7	258	4	US-09-252-991A-22452	Sequence 22452, A
c 20	149.5	9.7	511	4	US-09-252-991A-26078	Sequence 26078, A
c 21	148.5	9.6	190	4	US-09-252-991A-25320	Sequence 25320, A
c 22	148.5	9.6	300	4	US-09-252-991A-27679	Sequence 27679, A
c 23	148.5	9.6	467	4	US-09-252-991A-18296	Sequence 18296, A
c 24	148	9.6	335	4	US-09-252-991A-24046	Sequence 24046, A
c 25	148	9.8	375	4	US-09-252-991A-31128	Sequence 31128, A
c 26	148	9.6	710	4	US-09-252-991A-24946	Sequence 24946, A
c 27	147.5	9.5	341	4	US-09-252-991A-23424	Sequence 23424, A
c 28	147.5	9.5	568	4	US-09-252-991A-23264	Sequence 23264, A
c 29	146	9.4	316	4	US-09-252-991A-25345	Sequence 25345, A
c 30	146	9.6	335	4	US-09-252-991A-23948	Sequence 23948, A
c 31	145	9.4	187	4	US-09-252-991A-19743	Sequence 19743, A
c 32	144.5	9.5	316	4	US-09-252-991A-22720	Sequence 22720, A
c 33	144.5	9.5	763	4	US-09-252-991A-30146	Sequence 30146, A
c 34	143	9.2	200	4	US-09-252-991A-21290	Sequence 21290, A
c 35	143	9.4	441	4	US-09-252-991A-24396	Sequence 24396, A
c 36	142.5	9.2	335	4	US-09-252-991A-23674	Sequence 23674, A
c 37	142	9.2	228	4	US-09-252-991A-28381	Sequence 28381, A
c 38	142	9.2	376	4	US-09-252-991A-23577	Sequence 23577, A
c 39	142	9.2	398	4	US-09-252-991A-26217	Sequence 26217, A
c 40	142	9.4	441	4	US-09-252-991A-28965	Sequence 28965, A
c 41	141.5	9.1	228	4	US-09-252-991A-28783	Sequence 28783, A
c 42	141.5	9.1	239	4	US-09-252-991A-25387	Sequence 25387, A
c 43	141.5	9.1	325	4	US-09-252-991A-26580	Sequence 26580, A
c 44	141.5	9.1	369	4	US-09-252-991A-25394	Sequence 25394, A
c 45	141.5	9.1	433	4	US-09-252-991A-28695	Sequence 28695, A

ALIGNMENTS

RESULT 1	US-09-536-647-2	Application US/09536647
Sequence 2, Appl1	Patent No. 6579708	GENERAL INFORMATION:
APPLICANT: Johnson, Randall	APPLICANT: Ho, Yen Sen	TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
FILE REFERENCE: GP50020	CURRENT APPLICATION NUMBER: US/09/536,647	CURRENT FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 3	SOFTWARE: FastSeq for Windows Version 3.0	SEQ ID NO 2
LENGTH: 260	TYPE: PRT	ORGANISM: Human
US-09-536-647-2	US-09-536-647-2	
Alignment Scores:		
Pred. No.:	1.64e-130	Length: 260
Score:	1363.00	Matches: 260
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	90.03%	Indels: 0
DB:	4	Gaps: 0
US-09-896-522-3 (1-834) x US-09-536-647-2 (1-260)		
QY	1	ATGGCTTGGCGGGGCGAAGACTGCGAGCCCGGAGGCGAGCGTCCGCAC 60
DB	1	MetAlaSerAlaGlyGlyGlnAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
QY	61	CAGCGGCGCTTCCTGATAGAGGTGAGCGGCGGCGACTGCGAGCGGAGAGTGCACCGTGTG 120
DB	21	GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40

```
QY      121 GAGAAATCATGAGTTGCTGGGACAGAACGAGGTGGACAGCCGGACCGGAAGTGTGTC 180
      |||
Db      41 GlnlylelmetcgluLeuLeuGlInasnglIuValGluglnArgGlnArgValVal 60
QY      181 ATCTGAGCCGAGCAGGATTTCTACAAGGTCTCTGACGGCAGAGAGCAAGGCGCTTG 240
      |||
Db      61 IleuSerGlnmaphrpheryrlyValValLeuThrAlaGlnGlnlySalAluValLeu 80
QY      241 AAAGACAGTACAAATTTTGACATCCAGATGCTTTTGATTAATGATTTGATGACAGAGCT 300
      |||
Db      81 LysGlyGlnTyraenPheasphIsProAspAlaPheasphasndPheumethIsArgThr 100
QY      301 CTGAAGAAATCTGTGGAGGGCAAAACGGTGGAGGTGGCGGACCTTATGATTTTGGACAC 360
      |||
Db      101 LeuLysasnIleValGluglyLysThrValGlIuValProThrTyraAspPheValThrHis 120
QY      361 TCAAGGTTTACAGAGACCAAGGTGTCTACCTCGGACGAGTGTCTGTGTTGAGGCGATC 420
      |||
Db      121 SerArgLeuProGlnThrThrValValTyrrProAlaAspValValLeuPheGlnGlyIle 140
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
      |||
Db      141 LeuValPheTyrrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY      481 GACTCCGACGTCAAGCTGTCTCGAAAGATTCTCGGGACGTCGGCGGAGGAGGACCTG 540
      |||
Db      161 AspSerAspValAlaGlnLeuSerArgValLeuArgAspValAlaGlnArgGlyArgAspLeu 180
QY      541 GAGCAGATTCTGACGACATACACACTTCGTGAAGCGGCTTCGAGAGATTCTGCTG 600
      |||
Db      181 GlnGlnIleLeuThrGlnTyrrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY      601 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 660
      |||
Db      201 ProThrLysLysTyrrAlaAspValIleIleProArgGlyValaAspAsnMetValAlaIle 220
QY      661 AACCTGATTCGCGACATCCAGACATTTGTAATGTGACATCTGCAATATGCGACCGA 720
      |||
Db      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysPrlCysLysrTPHIsArg 240
QY      721 GAGAGGTCCAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGAGGAGCAACCTG 780
      |||
Db      241 GlnGlySerAsnGlnArgSerTyrrLysArgThrPheSerGlnProGlnLysPheHisProGly 260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:
Pred. No.: 1,27e-126 Length: 277
Score: 1325.00 Matches: 254
Percent Similarity: 95.67% Conservative: 11
Best Local Similarity: 91.70% Mismatches: 12
Query Match: 87.52% Indels: 0
Gaps: 0

US-09-896-522-3 (1-834) x US-09-536-647-3 (1-277)
```

```
QY      1 ATGGCTTCCGCGGAGGACCGAAGACTGCGAGAGCCCGCGGAGGCGGACCGTCCGCAC 60
      |||
Db      1 MetAlaSerAlaGlyGlyGlySerGlnSerAlaAlaProGlnLysAspArgProGln 20
QY      61 CAGCGGCGCTCTCTGTAATAGGAGGTGACGGCGGCGACTCCAGCGGAGAACTGACCTGTGT 120
      |||
Db      21 ProArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY      121 GAGAAATCATGAGTTGCTGGGACAGAACGAGGTGGACAGCCGGACCGGAAGTGTGTC 180
      |||
Db      41 GlnlylelIleMetcgluLeuLeuGlnAsnGlnIuValaAspArgArgGlnArgLysLeuVal 60
QY      181 ATCTGAGCCGAGCAGGATTTCTACAAGGTCTCTGACGGCAGAGAGCAAGGCGCTTG 240
      |||
Db      61 IleuSerGlnmaphrpheryrlyValValLeuThrAlaGlnGlnlySalAluValLeu 80
QY      241 AAAGACAGTACAAATTTTGACATCCAGATGCTTTTGATTAATGATTTGATGACAGAGCT 300
      |||
Db      81 LysGlyGlnTyraenPheasphIsProAspAlaPheasphasndPheumethIsArgThr 100
QY      301 CTGAAGAAATCTGTGGAGGGCAAAACGGTGGAGGTGGCGGACCTTATGATTTTGGACAC 360
      |||
Db      101 LeuLysasnIleValGluglyLysThrValGlIuValProThrTyraAspPheValThrHis 120
QY      361 TCAAGGTTTACAGAGACCAAGGTGTCTACCTCGGACGAGTGTCTGTGTTGAGGCGATC 420
      |||
Db      121 SerArgLeuProGlnThrThrValValTyrrProAlaAspValValLeuPheGlnGlyIle 140
QY      421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCTGTGACACC 480
      |||
Db      141 LeuValPheTyrrThrGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY      481 GACTCCGACGTCAAGCTGTCTCGAAAGATTCTCGGGACGTCGGCGGAGGAGGACCTG 540
      |||
Db      161 AspSerAspValAlaGlnLeuSerArgValLeuArgAspValGlnArgGlyArgAspLeu 180
QY      541 GAGCAGATTCTGACGACATACACACTTCGTGAAGCGGCTTCGAGAGATTCTGCTG 600
      |||
Db      181 GlnGlnIleLeuThrGlnTyrrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
QY      601 CCGACAAAGAAGTATGCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCCATC 660
      |||
Db      201 ProThrLysLysTyrrAlaAspValIleIleProArgGlyValaAspAsnMetValAlaIle 220
QY      661 AACCTGATTCGCGACATCCAGACATTTGTAATGTGACATCTGCAATATGCGACCGA 720
      |||
Db      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysPheCysLysArgHisArg 240
QY      721 GAGAGGTCCAATGGGCGGAGCTCAAGCGGACCTTTCTGAGCGAGGAGCAACCTG 780
      |||
Db      241 GlnGlyProAsnGlnLysArgSerTyrrLysArgThrPheProGlnLysPheHisProGly 260
QY      781 ATCTGACCTCTGGCAAGCGTCAATTTGATTCGAGAGGAGCAAGACCCGAC 831
      |||
Db      261 ValLeuAlaThrGlnLysArgSerHisLeuGlnLysSerSerArgProHis 277

RESULT 3
US-09-134-001C-3618
; Sequence 3618, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucetle-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
```

```
/ SEQ ID NO 3618
/ LENGTH: 216
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3618

Alignment Scores:
Pred. No.: 1,49e-37 Length: 216
Score: 451.00 Matches: 91
Percent Similarity: 63.68% Conservative: 44
Best Local Similarity: 42.92% Mismatches: 61
Query Match: 29.79% Indels: 16
Gaps: 4

US-09-896-522-3 (1-834) x US-09-134-001C-3618 (1-216)

QY 73 CTGATAGGGGTGACGGCGGCACTGCCAGGGAAGTCGACCGTGTGGAAGATCATG 132
Db 15 lletlelglylealaglyserglyserglyserthrvalthrsealalelmet 34
QY 133 GAGTTGCTGGGACAGAAACGAGTGGAAACAGCGGACGGGAGAGTGGTCATCTGAGCCAG 192
Db 35 Lysasmlen-----GluclyhisserValalauleuAlaGln 47
QY 193 GACAGGTTCTACAAG-----GTCTGACGCGACAGACGAGAAGGCCAGCCTTGAA 243
Db 48 AsptlyrtyrtyrlysaerplnserhisbleuthrPheglucluargleuGluThr----- 65
QY 244 GACAGTACAATTTTGACATCCAGATGCTTGTATATGATTTGATGACACAGACTTG 303
Db 66 -----AsnlyrasphisproPhealaPheaspsasmlenleuilehisasmlen 82
QY 304 AAGAATCGTGGAGGGCAAAACGCTGAGAGTGGCCGACCTATGATTTTGTGACACTCA 363
Db 83 LysAspleuarglsnnglylvsProValGluValProthrlyraspyrserGlnhisThr 102
QY 364 AGGTTACGACGACGACCGGTGTCTACCCCTGCGACGCTGCTGTTGAGGCACTTG 423
Db 103 ArgSerlysgluThrlealaPheasprolysaPvalalleleValGluGlylePne 122
QY 424 GTGTTCTACGACGACGACATCGGACATGTTCCACCTGCGCTTCTGTGACACCGAC 483
Db 123 AlaIeuglunsnanThrleuArgasPmetMetAspVallyserlyValAspThrAsp 142
QY 484 TCCGAGTCAAGCTGTCTCGAAGATTTCCGCGGACGTCGC--CGAGGAGGAGACTG 540
Db 143 AlaAspleuarglleuArgleuThrArgAspThrlysgluArgGlyArgThrMet 162
QY 541 GACGACATTTCTGACGACATCACACCTTGTGAGCGCGCTTCCGAGAGTTCTGCTG 600
Db 163 GluSerValilleasnglnlyrleuasnaValaIargProMetHisgluGlnPheleGlu 182
QY 601 CCGACAAGAAGATGATCGATGATCATCCACGAGAGTGGACATATGATGTCATC 660
Db 183 ProThrlyserHisAlaAspIlelleleProGluGlylyserHisValAlaIle 202
QY 661 AACCTGATCGTGCAGACATCCGAGACATTTGAT 696
Db 203 AspIleMetThrThrlyserIleGlnSerLeuValSer 214

RESULT 4
US-09-107-532A-5024
/ Sequence 5024, Application US/09107532A
/ Patent No. 6585275
/ GENERAL INFORMATION:
/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
```

```
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD/ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 5024:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 210 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHEICAL TYPE: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (B) LOCATION 1...210
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
US-09-107-532A-5024

Alignment Scores:
Pred. No.: 7.95e-36 Length: 210
Score: 434.00 Matches: 83
Percent Similarity: 62.33% Conservative: 51
Best Local Similarity: 38.60% Mismatches: 65
Query Match: 28.67% Indels: 16
Gaps: 4

US-09-896-522-3 (1-834) x US-09-107-532A-5024 (1-210)

QY 64 CGGCCCTTCTGATAGGGGTGAGCGGCGGCACTGCCAGCGGGAAGTGCACCGTGTGAG 123
Db 6 LysProIlellellelyvalthrlyglyserglyserglyserthrvalserArg 25
QY 124 AAGATCATGAGATTTGCTGGACAGAAACGAGTGGAAACGCGACGCGAAGTGTGATC 183
Db 26 AlaIlePheasAsnThrProAspHis-----SerIleMetMet 38
QY 184 CTGAGCCAGACAGGTTCTTACAAG-----GTCTGACGCGACAGACGAGAAGCCCAAG 234
Db 39 LeuGluGlnAspSerlyrtyrlysaerplnserhisbleuSerPheglucluargleuAsn 58
QY 235 GCCTTGAAAGACAGATGATTTTGACCATCCAGATGCTTGTATATGATTTGATGCAC 294
Db 59 Thr-----AsnlyrasphisproPhealaPheaspsasmlenleuile 73
QY 295 AGGACTGGAAGACATCGTGGAGGCAAAAACGCTGAGAGTGGCCGACCTATGATTTTGTG 354
Db 74 GlnHisValGlyAspleuLeuasnlyrlysaIalleGluLyProVallyrAspyrVal 93
QY 355 AACACTCAAGTTACGACGACGCGGTGTCTACCCCTGCGGACGTCGTTGTTGAG 414
Db 94 AlaHisThrArgserGlnAlaThrIlelleGlnGluProlysgluValalleleGluGln 113
QY 415 GGCATCTTGTTCTTACAGCAGGAGATCCGCGACATGTTCCACCTGCGCTTTCGTG 474
```

Db 114 GlyIleuLleuGluAspGluArgLeuArgAspLeuMetAspIleValIleVal 133

Qy 475 GACACCGACCTCCGACGCTGCTCTCGAAGATTCTCCGGAGCTG---CGCCGAGG 531

Db 134 AspThrAspAspAspIleArgIleIleArgIleArgIleValSerGAspMetGluGluArgIly 153

Qy 532 AGGACCTCGACGACGATTTCTGACGAGTACACCACTTCCTCGAAGCCGCTTGAAGAG 591

Db 154 ArgThrLeuAspSerValIleGluGlnIleThrValValValSerProMetTyrHisGln 173

Qy 592 TTCTGCTCCGCGACAAAGAGTATGCCGATGATGATCATCCACGAGAGTGAGCAATATG 651

Db 174 PheIleGluProThrIleArgIleArgIleIleValProGluGlyGluAsnHis 193

Qy 652 GTTGCATCAACCTGATCGTCGACGACATCCAGACGACATTTGAT 696

Db 194 ValAlaIleAspLeuIleThrThrIleValAlaSerPheLeuAsn 208

RESULT 5

US-09-198-452A-783

Sequence 783, Application US/09198452A

Patent No. 6559294

GENERAL INFORMATION:

APPLICANT: Griffla, R.

TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments, and uses thereof, in particular for the diagnosis, prevention and treatment of infection

TITLE OF INVENTION: and treatment of infection

FILE REFERENCE: 9710-003-999

CURRENT APPLICATION NUMBER: US/09/198,452A

CURRENT FILING DATE: 1998-11-24

NUMBER OF SEQ ID NOS: 6849

SEQ ID NO 783

LENGTH: 222

TYPE: PRT

ORGANISM: Chlamydia pneumoniae

US-09-198-452A-783

Alignment Scores:

Pred. No.: 1,436-33 Length: 222

Score: 412.00 Matches: 84

Percent Similarity: 63.01% Conservative: 54

Best Local Similarity: 38.36% Mismatches: 59

Query Match: 27.21% Indels: 22

DB: 4 Gaps: 6

US-09-896-522-3 (1-834) x US-09-198-452A-783 (1-222)

Qy 73 CTGATAGGGGTGAGCGGGCGGCACTGCCAGCGGAGTGCACCTGTGTGAGAAATCATG 132

Db 8 IletIleGlyIleThrGlyIleSerGlyAlaGlyIleThrThrLeuThrGluAsnIleLys 27

Qy 133 GAGTTGCTGGGACAGACAGAGGTGAGACAGCGCGACGCGAAGGTGTCATCCTGAGCCAG 192

Db 28 GluIlePheGlyGluAsp-----ValSerValIleGln 39

Qy 193 GACAGGTTCTACAG-----GTCTGACGCGACAGACAGAACGCCAGCCTTGAAA 243

Db 40 AspAsnIleTyrIleLysAspArgSerHisTyrThrProGluGluArgAlaAsnLeuIle--- 58

Qy 244 GACACGTACAAATTTGACCATTCACATGCTTTGATATGATTGATGACACAGACTGTG 303

Db 59 -----TTPAspHisProAspAlaPheAspAsnAspLeuIleIleSerAspIle 74

Qy 304 AAGACATCGTGAGAGGCAAAACGTTGGAGGTGCGGACCTGATGATTGTGTGACACATCA 363

Db 75 LysArgLeuLysAsnAsnGluIleValGlnAlaProValPheAspPheValLeuGlyAsn 94

Qy 364 AGG---TTACACAGACACACGCTGTGCTTACCTTCGCGAGCGTGTCTGTTGAGGGCATC 420

Db 95 ArgSerLysThrGluIleGluThrIleTyrProSerLysValIleLeuValGluGlyIle 480

Qy 421 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCACCTCGCGCTTGTGTGACACC 480

Db 115 LeuVal]PhegiuaenGIngluueArAspleumetcrspleleArIlePheValAspThr 134
QY 481 GACTCCGACGTGACCTGTCTCGAAGAGTTCTCCGGACGTCGC--CGAGGAGGAC 537
Db 135 AsplaAeSpIuaArgIleLeuArGArMetValArAspValGIngluGIngluYAspSer 154
QY 538 CTGAGACGATTTCTGACGACGACTACACACCTCTCTGTAAGCCGGCTTCGAGGATTCGC 597
Db 155 ValAspCylleMetSerArgTylLeuSerMetValAspProMetHisGluYsPheIle 174
QY 598 CTGCCGACAAAGATATGACCGATGTATCATCCACGAGAGTGACCAATATGTTGCC 657
Db 175 GluProThrArgIlyrAlaAspIleIleValHisGlyenrYrArgGlnAenValVal 194
QY 658 ATCAACCTGATCGTCAG-----CACATCCAGACATCTGAAATGTGAC 702
Db 195 ThrAenIleuSerGInlyrIleYAsnHisIleuGlnAenAlaLeuGlnuSerAsp 213

RESULT 6
US-09-252-991A-24671
/ Sequence 24671, Application US/09252991A
/ Patent No. 6551795
GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 24671
/ LENGTH: 1150
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24671

Alignment Scores:
Pred. No.: 64e-08 Length: 1150
Score: 164.00 Matches: 70
Percent Similarity: 31.06% Conservative: 12
Best Local Similarity: 26.52% Mismatch: 108
Query Match: 10.60% Indels: 74
DB: 4 Gaps: 13

US-09-896-522-3 (1-834) x US-09-252-991A-24671 (1-1150)
QY 709 TGCAGATGTCAACATTCAGAAATGCTCGATGTCTGCACAGATCAGGTTGATGGCAACA 650
Db 13 CysAlaThrArgAenAlaSerCysCysAlaCysThrArg-----TrrArgPro 28
QY 649 TATGTCCACTCTCTGGGAGATGATACATCGGACATCTTTGTGGCAGGACGAAC 590
Db 29 ValAlaPro----- 31
QY 589 CCTGGAAGCCGGGCTTACGAGAGTGGTACTCGGTCAGAAATGTGCTCCAGATCCCTCC 530
Db 32 -----ArgArgTrpSerSerAlaArgThrSerAlaProCysProIys 45
QY 529 CTCGGCGACGT-----CCCGAGAACTTTTGAGACA 497
Db 46 AlaCysAlaAaArgAenTrpHisAlaAlaIysCysThrValProAlaCysCysSerSerSer 65
QY 496 GCGTGAAGCTGCGAAGCTGGTGTCCAGAGAGCGCCAGAGTGACA----- 452
Db 66 GluAlaAlaAenAlaSerCysProAlaArgAenGlyrArgArgTrpProTrpProProPro 85
QY 451 -----TGTCCGGAGATCTCTGTGCTTAAACA-----CCA 422

```

Db      86 AlAProThrThrcysargAlaCysProAlaGlyGlyAlaIaArgThrSerThrglyAla 105
Qy      421 AGATGCCCTGAAACAGAACACCGTCGCGAGGGTGAACCGGTGTCTGTGTAACCTTG 362
Db      106 AlAAlaProAspThrAlaAlaProAlaArgAsnAlaArgProThrThrPaaenCysAlaProArg 125
Qy      361 AGTGTG-----TCACAAATATCATAGTCGGGACCTCCACCGTTTGGCCCT 317
Db      126 SerAlaGlyProThrProProAlaSerAsnArgSerIleProThrPro-----Pro 143
Qy      316 CCA-----CGATGTTCTTCAAGAGTCTGTGCATCAATCATATTAACAAGGCATGTGAT 263
Db      144 ProArgSerArgThrProAlaSerAlaMetSerGlyGlnProProArgArgTyr--- 162
Qy      262 GGTCAAAATGTACTGTCCTTCAAGGCGTTGGCTTCTGTCTGTCCCGTCAAGACCTTGT 203
Db      163 -----CysTyrCys-----ProTyrArgThrProThrAlaAlaTyrPseSer 176
Qy      202 AGAACCTGTCTGCTGAGATGACACCACTTCGCGCTCCGCTGTTCACCTCGTTCTGTG 143
Db      177 ArgThrCysProAlaProAlaThrProThrArgThrGlyGlyAlaGlyAlaGlyTyr 196
Qy      142 CCAGCAATCCATGATCTTCTCACACACGTCGACTTCGCCGTCGACGTCGCCGCTCA 83
Db      197 ProArgArgCys---SerGluArgArgArgSerThrGlyAlaCysAlaCys----- 212
Qy      82 CCCTTACAGAGAGGCGCGTGTGCGGACGTCGCGCTCCGCGCGGGGCTCTGCG--- 26
Db      213 -----SerAsnAlaAlaAlaGlyMetProArgArgProMetIleGluProArgAla 230
Qy      25 AGTCTTCGCTC 14
Db      231 ThrLeuArgLeu 234

```

```

RESULT 7
US-09-199-637A-369
; Sequence 369, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-199-637A-369

```

```

Alignment Scores:
Pred. No.:      7.79e-08      Length:      265
Score:          160.00      Matches:      87
Percent Similarity: 35.62%      Conserves: 22
Best Local Similarity: 28.43%      Mismatches: 107
Query Match:     10.34%      Indels:      90
Db:              4          Gaps:      17

```

```

US-09-896-522-3 (1-834) x US-09-199-637A-369 (1-265)
Qy      820 TGCTGAGCTCCAAATGTAGCCGTTTGGCCAGAGGTACGATCCAGGAGGTGTCCCTGAGCT 761

```

```

Db      11 CysArgSerProAsnValProArgCysCysArgSerGlyCys---GlyArgArgAlaAla 29
Qy      760 CAGAAAAGTCCGCT-----TGAGTCCGCCAT----- 731
Db      30 GluSerArgSerIleProThrProAlaProCysAlaAlaGlyAsnProProArgThrCys 49
Qy      730 -----TGACCCCTCCGCGTGCATTTGGAGTGCACCATTCACAAATGCTGGATGT 677
Db      50 ArgGlySerSerAspLeuSerProValProArgAlaValProGlyGlnAlaAlaSerThr 69
Qy      676 GCTGACGATCAGAGTTGATGACCAACCATATTTGCCATCTCTCGTGGAGTATCATCCG 617
Db      70 ThrArgArgSerArgArgLeuProProAlaAlaProCys-----Arg 83
Qy      616 CATACTTCTTTTGGCAGAGGACGAACCTCTCGAAGCCGCGCTTACGAAAGTGTACT 557
Db      84 HisGlySerArgHisAla---ArgSerProArgArgProLeuAspArgArg-----Arg 100
Qy      556 GCGTCAGAAATCGCTCCAGGTCCTCCGCGGCGACGTCGCGGAGAACTCTTCGAGACA 497
Db      101 AlaThrSerArgAlaPro---ProSerValHisAla----- 111
Qy      496 GCGTCAGTCCGAGTCCGTGTCCACGAAAGGCGCAGTGGAAACATGTCGCGATCTCT 437
Db      112 -----ThrArgAspTyrPargIleValAlaGlyMetProAlaMetSerAlaPro 127
Qy      436 GCGCTTGAAACACCAAGATGCCCTCAACAGAAACACGTCGCGAGGTGACCACTGTG 377
Db      128 Gly-----GluAspCysAlaHisTyr---ProArgProThrCysArgProPro--- 142
Qy      376 TCTCTGTAACTTGAGTGTGTACAAAAATCATAGTGGGACCTCCACCGTTTCCCT 317
Db      143 -----AlaGlySerLeuArgProAlaArgAlaProProArgArgPro 156
Qy      316 CCAAGATGTTCTTCAAGATGCTGTGCATCAATATCAAAAGCATGTGATGTCAA 257
Db      157 ProArgSerAlaLeuProSerThrAla-----GlyCys 167
Qy      256 AATTGACTGTCTTTCAAAGGCTTGCGCTTCTGTCTGCGC-----TAGGA 209
Db      168 GlyCysVal-----ProThrAlaValProArgSerGlyArgProGly 181
Qy      208 CCTGTGAAAC----- 197
Db      182 ProAlaSerThrGlyGlnArgTyrGlyProArgAspArgAlaArgSerArgAlaAsn 201
Qy      196 ---TGCTGAGTCAAGATGACCACTTCGCGTCCGCGTGTTCACCTGTTCTGCCA 140
Db      202 AlAAlaProGlyProGlyArgSerProAlaThrSerAla---ProGlyArgSerProPro 220
Qy      139 GCAACTCCATGATCTTTCACACACGCTGACTTCCGCTGGCGATGCGCGCTCACCC 80
Db      221 GlyCys-----SerAlaThrAspArgAlaAlaArgArgSerArg 233
Qy      79 CTATCAGAAAGGCGCGTGTGCGACGTCGCGCTCCGCGCGGCGCTCTCGCAATCTT 20
Db      234 LeuProProArgMetProHisAlaHisSerArgProAlaProProGlySerProAlaVal 253
Qy      19 CGCCTCCCGCCGAGACCA 2
Db      254 ArgArgAlaProAlaPro 259

```

```

RESULT 8
US-09-252-991A-25569
; Sequence 25569, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25569
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25569

Alignment Scores:
Pred. No.: 1.13e-07 Length: 274
Score: 158.50 Matches: 72
Percent Similarity: 35.50% Conservative: 21
Best Local Similarity: 27.48% Mismatches: 78
Query Match: 10.25% Indels: 91
DB: Gaps: 16

US-09-896-522-3 (1-834) x US-09-252-991A-25569 (1-274)
QY 763 GCTCAGAAAAGCTCCGCTTGT-----AGCTCCGCCATTGGACCTTCCTCGTGCATT 710
Db 88 AlaArgArgArgSerSerValArgArgAsnArgGlyHisThrArgValAlaThrAla 107
QY 709 TGCAGATGTCACTTCAGATGTCCT-----GGATGCTGTCAACATCAGTTGA 659
Db 108 ProArgCysArgArgArgArgGlyThrProAlaAlaSerGlyAsnProAlaArgAla----- 125
QY 658 TGGCAACCATATTTGTCACCTCCTCGTGAGATGATCATCATCGCATACTTCTTTGCGCA 599
Db 126 -----ArgHisThrAla----- 129
QY 598 GCGAAGAACTCTCGAAGCGCGCTTTCAGAAAGTGTTGACT---GCGTCAGAAATGCT 542
Db 130 -----ProArgArgArgAlaArgArgGlyCysAlaArgAlaSerArgGly 146
QY 541 CCAAGTCCCTCCCTCGGGCGACGTCGCGAAGAACTCTTCAGAGACGCTGACGTCGAGT 482
Db 147 ProGlyGly----- 149
QY 481 CGGTCGCAAGAGAGCGCAGGTGA-----ACATGTCGCGATCTCTGCTGTAGA 428
Db 150 -----ProArgArgAlaAlaArgAlaProValAlaAlaProGlySerProGly--Arg 166
QY 427 ACACCAAGATGCTCTCAACAGAACACACGTCGCGAGGTAGACCAACGCTGCTTGTGTA 368
Db 167 SerAlaThrAlaArgProSerAlaProArgAlaProAlaThrProPro----- 182
QY 367 ACCTTGAGTGTGTACAAATCATAGTCCGACCTCCACCGTTTTCCTCCACATGT 308
Db 183 -----AspAlaProProAlaArgProAlaAlaCys 192
QY 307 TCTTCAGAGTCTGTGATCAATCATTAATCAAGATCGATGTGATGATCAAAATGTACT 248
Db 193 AlaAlaSerGlyCysArgSerTrpArgTrpArgTrpArgTrpArgTrpArgTrpAlaVal 212
QY 247 GTCTCTCAAGGCTTGCGCTTCTGCTGCTGCGTACAGACCTGTGAACTGTCTGGC 188
Db 213 ArgAlaTyArgPro-----LeuSerArgProProCys--ProCysAlaGly 227
QY 187 TCGAGTAGACCACTTCGCTGCGCTGTTCACCTGTTCTGTCGACGAACCTCCATGA 128
Db 228 AlaGly-----SerAlaGlyAlaCysArgSerArg--LeuPro----- 239
QY 127 TCTTTCACACACGATGACTTCCGCTGCGACGTGCGC---CGTACCCCTATCAGA 71
Db 240 -----ArgCysArgProGlySerProSerSerGly 249
QY 70 AGGCGCCTGAGTGCAGACGCTCGGCTCGCGCGGCGCTCTCGCAGTCTTCGCTCGG 11
Db 250 ArgSerCys-----ArgProProAlaSerProGlyGluSerProCysIlePro 265

```

```

QY 10 CCGAG 5
Db 266 AlaArg 267

RESULT 9
US-09-252-991A-26899
; Sequence 26899: Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26899
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26899

Alignment Scores:
Pred. No.: 1.35e-07 Length: 247
Score: 157.50 Matches: 83
Percent Similarity: 34.60% Conservative: 17
Best Local Similarity: 28.72% Mismatches: 84
Query Match: 10.18% Indels: 105
DB: Gaps: 21

US-09-896-522-3 (1-834) x US-09-252-991A-26899 (1-247)
QY 688 TGTTCAGATGTCGTCGACGATCAGATGATGCAACATATGTCGACTCCTCGTGGGA 629
Db 2 CysProSerCysSerProProSer-----ProCysSerProLeuAlaArgTrp 17
QY 628 TGATCATGCGCATCTTCTTTGCGGACGAGCACTCT-----CGAAG 581
Db 18 ProSerProThrProAlaThrAlaAsnThrAlaTrpSerProAlaSerProIleArgSer 37
QY 580 CCGGCTTCACGAAGGTGCTGACTGCGTCAAGATTCGCTCCAGGTCCTTCCTCGCGCA 521
Db 38 ProAlaSerThrThrCysTrpArgCysSerProSerAlaSerGlyProProSerAsnArg 57
QY 520 CGTCCGAGAACTCTTCGACAGCGCTGACGTCGAGTCCGGTGCACGAAGA----- 467
Db 58 GluPro-----ArgValSerProCysProAlaProSerSer 69
QY 466 -----GGCGACAGTGAACATGTCGCGATCTCTGCTGTAGAAC----- 425
Db 70 AlaProCysTrpSerAlaGlyCysSerAlaSerArgAlaCysAsnCysArgSerTrpIle 89
QY 424 -----CCAAGATGC----- 416
Db 90 ProAlaSerProProArgCysSerProSerAlaSerAlaTrpProSerProCysAlaArg 109
QY 415 -----CCTCAAGACGACGTCGCGAGGTAGACA----- 383
Db 110 ArgCysArgTrpArgTrpProProProArgCysSerProTrpProThrGlySerProThr 129
QY 382 CCGTGTCTCTGTGTAACCTTGAAGTGTGCAAAATCATAGGTGCGACCTCCACCGTTT 323
Db 130 AlaTrpSerCys-----ArgThrSerProAlaProGly 140
QY 322 TGC-----CCTCAAGATGTTCTTCAGATGC----- 296
Db 141 CysThrArgSerValSerSerProProProProArgCysThrProProAlaThrArgTrp 160

```

```

QY 295 TGTGATCAATATCATTAATCAAGGATGTGGATGTGCAAAATTGTACTGTCCTTTCAAG 236
    |||||
Db 161 CysAlaSer-----TTPAArgArg-----ProArg 169
QY 235 CCTGGCCTTCTGCTCTGCGTCAGACCTTGAGAACCTGCTCTGCGTCAGAGTGA 176
    |||||
Db 170 ProTyr--TyrValSerProGlyTyrPro--ArgArgAlaProGlySerGlyCysTyr 187
QY 175 CCTTCGCGCTGCGCTGTTCCACCTCTGCTCTGCTCCAGAACCTCATATCTTCTCACACA 116
    |||||
Db 188 Pro-----AlaGluProProArg--LeuPro-----GlnArgLeu 198
QY 115 CGGTGCACTTCCGCGC--TGGCAGTCCGCGCGCTCACCC--CTATCAGAAAGGCGCT 62
    |||||
Db 199 TlSerThrArgArgProTyrCysThrArgArgArgProTyrAlaCysAlaSerIleAla 218
QY 61 GGTGCGGACGCTCGG-----CCTCGCGCGCGG-----35
Db 219 ArgAlaSerGlyArgSerSerThrGlyValProGluThrArgSerAlaSerSerAsp 238
QY 34 GCGTCTCGCAGTCTTCCCTCCGCGCGC 8
    |||||
Db 239 GlySerHisSerMetLeuSerProPro 247

RESULT 10
US-09-252-991A-31116
/ Sequence 31116, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 31116
/ LENGTH: 664
/ TYPE: PRN
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31116

Alignment Scores:
Pred. No.: 2,51e-07 Length: 664
Score: 157.00 Matches: 90
Percent Similarity: 29.09% Conservative: 15
Beet Local Similarity: 24.93% Mismatches: 105
Query Match: 10.37% Indels: 151
DB: 4 Gaps: 16

US-09-896-522-3 (1-834) x US-09-252-991A-31116 (1-664)
QY 15 AGCGGAAGACTGCGAGAGCCC-----CGCGCGGAGGCGGACCGCTCCGACACGCG 65
    |||||
Db 265 ArgArgArgArgAlaSerProTyrProArgArgArgArgGlyArgProGlyGlnGlyAla 284
QY 66 GCGCTTCTATAGGGGTGAGCGGCGGCACTGCGACGAGGGAAGTCGACCGCTGTGTGAA 125
    |||||
Db 285 Ala-----HisArgArgGlyHisArgGlyArgProArgSerGlyAla 298
QY 126 GATCATGAGATT-----GCTGGACAGAACGAGGTGGAACAGCG 164
    |||||
Db 299 ArgHisGlyValHisAlaSerArgArgGluGlnProGlySerGlyProGlyArgPro 318
QY 165 GCGAGCGAAGGTGTCTCTGAGCCGAGACAGGTTTACAAGT-----209
    |||||
Db 319 AlaProGlyArgProArgProAlaGlnGlyGlyGlnGlyArgHisArgLeuArg 338
QY 210 -----CCTGACGCGACAGACGAAAGCGCCTTGAAGGACATGATTTTGA 260

```

```

Db 339 ProHisArgProGlyGlyGlnGlyArgGlyHisGly-----350
QY 261 CCAATCCAGATGCTTTGATTAATGATTTGATGCAACAGAACTGTGAAGACATGTGGAGG 320
    |||||
Db 351 -----GlnGluLupProGlyGly 356
QY 321 CAAAAGGTGAGAGGTGCGACCTATGATTTTGTGACACATCAAGTTACAGAGAC 380
    |||||
Db 357 ArgArgGlyAlaGly-----AspArg 363
QY 381 GGTGTCTACCTCTGCGGACGTGCTGTGTAAGG-----416
    |||||
Db 364 GlyAlaLeuProValProArgGlnHisProArgGlyValAspHisArgHisHisPro 383
QY 417 -----CATGTT-----CGTGTTCACGCCAGGA 440
    |||||
Db 384 AlaValAspAlaValHisSerHisArgAspGlyArgGlnGlnGlyGlnArgGlnProAsp 403
QY 441 GATCCGGACATGTTCCACCTGCGGCTCTGCGACACCCGACCTCGGACGTGAGCTGTC 500
    |||||
Db 404 GluProArgArg-----AlaGlyLeuArgHisHisArgArgArgArgGlyAsp 420
QY 501 TCGAAGAGT--TCTCCGAGACGTGCGCGGAGGAGGGA-----536
    |||||
Db 421 ArgArgGluArgHisProProProGlyProArgProGlyProProArgProProAlaAsp 440
QY 537 -----CCTGAGCGAGAT 548
    |||||
Db 441 ProArgGlyThrProArgGlyLeuArgArgValGlnGlyGlyAlaProGlyAlaGly 460
QY 549 TCGAGGAGTACACACCTTGTGTAAGCGCGCTTCGAGAGTTGCTGCTGCGGACAA 608
    |||||
Db 461 LeuArgProAspHis-----HisGlyGlyValProAlaAspLeu 474
QY 609 GAAGTATGCCGATGTATCATCCACGAGAGTGCACATAT-----GGTTGCCAT 659
    |||||
Db 475 Arg-----ProHisArgGlyArgGlyGlyAspValProProAspGlyValHis 490
QY 660 CAACSTGATGTCGACGACACATCCAGACGACATTTGAATGTGCATC-----705
    |||||
Db 491 ArgGlyHisArgAlaAlaArg--ArgAspAspProLeuGlyAspLeuArgProGlyGlyAs 510
QY 706 -----TGCAAATGCAACCGAGAGGTCCAATGGCGGCACTAC-----744
    |||||
Db 510 ArgAlaValHisHisArgGlnGlyGlnGlyGlyGlyLeuArgHisAlaProArgAl 530
QY 745 -----AA 746
    |||||
Db 530 AspGlyLeuArgThrGlyAlaAlaLeuGlyAlaArgAlaProArgAlaGlyGlyArg 550
QY 747 GCGGACCTTTTTCGACGCGAGGAGCCACCTGGAGATGTCGACCTTGGCAACGATCACA 806
    |||||
Db 550 gATgProArgArgAspProAlHisArgAlaGlyGlyLeuAlaAspGlyGlnArgValHis 570
QY 807 T 807
    |||||
Db 570 s 570

RESULT 11
US-09-252-991A-24899
/ Sequence 24899, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190

```



```

QY 502 GAGACAGCTGACGCTCGAGTGGTGTTCACGAAAGGCGAGGTGGAACATGTCGCCGA 443
    ||||| ||| :||| ||||| ||||| |||||
Db 122 AlahrhlaahargahlaethrhclYArgAlaaglYarglYalaglYlY-----GlyGly 139
QY 442 TTCCTCGCTGTGAACACACAGATGCCCTCAACAGAACACGTCGCGAGGTAGACCA 383
    ||||| ||||| ||||| ||||| ||||| |||||
Db 140 ArgCysglYglYargYargProAlaThrProAspLeuThsleuLeuProGlyAlaAsp 159
QY 382 CCGGCTGCTCGTGAACCTTAGTGTGTCAACAAATCATAGTGTGGCA-----CCT 332
    ||||| ||||| ||||| ||||| ||||| |||||
Db 160 ProArgglYthrGlyglYalAaAspProAlaArgGlyleuArgProAlaThsArgGlyAsp 179
QY 331 CCACCGTTTGGCCCTCCACGATGTTCTTCAGAGTCTGTGATCAAAATCATATCAAGG 272
    ||||| ||||| ||||| ||||| ||||| |||||
Db 180 ArgProArgLeuProAlaArgY-----AlaGlnArgHisArgProThr 193
QY 271 CATGTGATGCTCAAAAATTGACTGTCTTCAAGGCTTGCCCTTGTCTGCTCCGTCA 212
    ||||| ||||| ||||| ||||| ||||| |||||
Db 194 HisArgAlaGlyGlnGlu-----GlnAspProArgGlyProAspSerLeuProGly 210
QY 211 GGACCTGTAGAACCTGTCTGTGCTCAGATGACACCTTCCGCTGCGCTTTCACACT 152
    ||||| ||||| ||||| ||||| ||||| |||||
Db 211 -----ThraAspProGly---GlyThrProArgAlaAlaGlyGlnArgProAla 225
QY 151 CGTTCTGCCAGCACTCCATGATCT-----TCTCAC 119
    ||||| ||||| ||||| ||||| ||||| |||||
Db 226 SerAspLeuProGlyValGlnArgGlyLeuLeuGlyLeuHisArgArgGlnProAspPro 245
QY 118 ACACGCTGACCTTCCGCTGACAGTGCAGCGCTCACCCTATCAGAAAGGCGCGTGT 59
    ||||| ||||| ||||| ||||| ||||| |||||
Db 246 ArgArgSerGlnArgArgGlyAspProArgArgProProAlaGlyArgAlaAlaAlaArg 265
QY 58 GCGACAGCTGGCGCTCCGGCGCGGGGCTCTGCAGTCTTGCGCTC 14
    ||||| ||||| ||||| ||||| ||||| |||||
Db 266 ProArgGlyProArgProAlaArgProAspValAlaAlaArgLeu 280

RESULT 13
US-09-252-991A-21436
; Sequence 21436, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 31142
; SEQ ID NO 21436
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21436

Alignment Scores:
Pred. No.: 2.24e-07 Length: 265
Score: 155.50 Matches: 84
Percent Similarity: 31.86% Conservative: 17
Best Local Similarity: 26.50% Mismatches: 99
Query Match: 10.05% Indels: 117
DB: 4 Gaps: 17

US-09-896-522-3 (1-834) x US-09-252-991A-21436 (1-265)
QY 775 GGATGATCCCGCTCGAGAAAGTCCGCTGTGAGTCCCGCCATTGGACCTCTCTCGGT 716
    ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GlyheProAlaProAlaArgSerAlaSerCysArgSerProAla----- 15
QY 715 GCATTTGCAGATGTCAATTCATTCAGATGTCCTGATGTGTGTGA-----CGATCAGGT 662

```

```

Db 16 ---ValProArgCysCysArgSerGlyCysGlyArgAlaAlaAlaLeuSerArgSerAla 34
    :||| ||||| ||||| ||||| ||||| |||||
QY 661 TGAATGCAACCAATATTGTCCATCTCTCGTGGATGATCAGATGGCATCTTCTTGTGCG 602
    ||||| ||||| ||||| ||||| ||||| |||||
Db 35 ProTrpProAlaProCys-----Ala 41
QY 601 GCAGGCAAGAACTCTCGAAGGCGCGGCTTCAAGAGGTGGTGTACTGTGTCAGAAATCTGT 542
    ||||| ||||| ||||| ||||| ||||| |||||
Db 42 AlaGlyAsnProProArgThr-----CysArgGlySer---SerAla 54
QY 541 CCAAGTCTCTCCCTCGGCGACGTCGCCGAGAACTTTCAGACAGACCTGACCTGGAGT 482
    ||||| ||||| ||||| ||||| ||||| |||||
Db 55 LeuSerProAlaProArgAlaValProGlyGlnAlaAlaSerThrThr---ArgArgSer 73
QY 481 CGGTGTCCA----- 473
    ||||| ||||| ||||| ||||| ||||| |||||
Db 74 ArgArgProProProAlaAlaProCysArgGlnGlySerArgHisAlaArgSerProArg 93
QY 472 -----CGAAGGCGCGAGGTGGAACATGTCCC----- 446
    ||||| ||||| ||||| ||||| ||||| |||||
Db 94 ArgAlaLeuAspArgArgArgAlaThrSerArgAlaProProSerAlaHisAlaValArg 113
QY 445 -----GGATCTCTGCTGAGTGAACACCA-----AGATGCCCTCAA 410
    ||||| ||||| ||||| ||||| ||||| |||||
Db 114 AspTrpArgGlnLeuAlaAlaGlyMetProAlaMetSerAlaProGlyGlnAspCysAlaHis 133
QY 409 ACAGAAACCACTCCGAGGATGACCAACCGTGTCTGTGTAACCTTGAATGTGTCAAA 350
    ||||| ||||| ||||| ||||| ||||| |||||
Db 134 Trp---ProArgProThrCysArgProPro-----AlaGlySer 145
QY 349 AATCATAGTGGCGACCTCCACCGTTTGGCTCCACAGATGTTCTTCAGAGTCTGTGCA 290
    ||||| ||||| ||||| ||||| ||||| |||||
Db 146 ArgArgProAlaArgAlaProProArgArgProProArgSerAlaLeuProSerThrAla 165
QY 289 TCAATCATATTCAAGGACATCTGATGTCAAAATTGATCTCTTCAAGGCTTGG 230
    ||||| ||||| ||||| ||||| ||||| |||||
Db 166 -----GlyCysAlaCysVal----- 170
QY 229 CTTTGTGCTGTGCG-----TCAGACCTTGTGAAC----- 197
    ||||| ||||| ||||| ||||| ||||| |||||
Db 171 ProThrAlaValProArgSerGlyArgProGlyProAlaSerThrGlyGlnArgGly 190
QY 196 -----TGTCTGCTCAGATATCACACT 173
    ||||| ||||| ||||| ||||| ||||| |||||
Db 191 ProArgAspArgAlaArgSerArgArgAlaAsnAlaAlaProGlyProGlyArgSerPro 210
QY 172 TCCGCTGCGCTGTTCACCTGCTGTGTCGCCAGCAATCCATGATCTTCTCACACAGG 113
    ||||| ||||| ||||| ||||| ||||| |||||
Db 211 AlaThrSerAla---ProGlyArgSerProProGlyCys----- 222
QY 112 TCGACTTCCCGCTGAGAGTGCAGCGCTCAACCCCTATCAGAAAGGCGCGTGTGCGGAC 53
    ||||| ||||| ||||| ||||| ||||| |||||
Db 223 SerAlaThrAspArgAlaAlaArgArgSerArgLeuProProArgGlyMetProHisAlaHis 242
QY 52 GGTGCGCTTCCGCGCGGCGCTTCCGAGTCTTCCGCTCCCGCGGAAGCA 2
    ||||| ||||| ||||| ||||| ||||| |||||
Db 243 SerArgProAlaProProGlySerProAlaValAlaArgArgAlaProAlaPro 259

RESULT 14
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```

/ PRIOR FILING DATE: 1998-07-27
 / NUMBER OF SEQ ID NOS: 33142
 / SEQ ID NO 30843
 / LENGTH: 663
 / TYPE: PRT
 / ORGANISM: Pseudomonas aeruginosa
 / US-09-252-991A-30843

Alignment Scores:

Pred. No.:	6,41e-07	Length:	663
Score:	153.00	Matches:	102
Percent Similarity:	27.88%	Conservative:	31
Best Local Similarity:	21.38%	Mismatches:	114
Query Match:	9.89%	Indels:	230
DB:	4	Gaps:	26

US-09-896-522-3 (1-834) x US-09-252-991A-30843 (1-663)

```

QY 832 AGTGGGCTCTGCTGTGACTCCAAATGTGACCGTTTGCAGAGGTGACATCCGAGGT 773
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 75 AlaGlyAlaCysCysTrpLysThrArg---AsnValAlaGlyArgSerAlaThrAspAla 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 ----- 773
DB 94 TrpSerSerAlaProAlaProAlaAsnCysCysLeuPheProAlaGlyProSerPro 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 772 -----GGTCCCGCTCAGAAA-----AGTCCGCTGT-----AGC 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 114 AlaSerProAlaProAlaAlaCysArgArgTrpProAlaGlyAlaAlaCysHisTrpProAla 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 739 TCCGCCATTGACCCCTCTCGTCCGATTTGC----- 707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 SerAlaTrpTrpLysLeuAlaProAlaArgCysCysTrpProAlaProProAlaProAla 153
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 707 ----- 707
DB 154 SerAlaGlyArgAlaCysCysAlaSerProAsnArgArgGluProTrpProPro 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 706 -----AGATGTCAACCATTCAGAAATGT-----CCTGATGT----- 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 174 SerProTrpAlaSerArgAlaGlyProAlaSerCysGlyArgProProAlaCysSerPro 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 -----GCTGCACGATCAGTTGTGGCAA 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 ValAlaThrAlaProThrAlaThrCysSerProProSerAlaArgSerAla---TrpLys 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 CCA---TATTGTCCACTCTCTGTGGAGATCAATCGCATCTTTCTTTCGCGAGGC 596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 ProCysAlaCysAlaLysAlaValGly-----SerAla 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 AGAATCTCTCGAAGG-----CCGAGCTCAGGAAGGTGTACTGCGTCAGAAATCTGCT 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 224 ArgSerProAlaAsnAlaTrpProAlaAlaSerAlaTrpCysProThrCysAsnSer--- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 CCAAGTCTCTCTCTCGCGCAGCTGCCGAGAACTTTCCAGACAGCCTGACGTGAGT 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 -----AlaArgProSerAlaIleAlaSerThrAla---ArgArgSer 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 -----CGTGTCCAGGAAGCGCAGGCGAAGCAATGTCTCC 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 ProSerThrAsnGlyArgProAlaAlaCysProThrThrThrProProAlaAsnAlaPro 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 445 GGATCTCTCTGCTGAGAACACCA----- 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 276 AlaSerAlaAlaAlaAsnTrpProTrpTrpLysAlaProSerProAlaMetProProSer 295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -----AGATGCCCTCAA 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 296 ThrSerAlaThrArgLeuAlaAlaCysGlyHisAspGlyAspAlaGlyArgAlaSerGln 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 409 ACAGAACCAAGCTCCG-----CAGCGTAGACCA 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 316 ThrProTrpArgAspThrSerProCysValProAsnCysValaAsnTrpArgArgArgIle 335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

QY 382 CCGNGTCTCTGGTAACCTTGAAGTGTGCACAAATCATGATGTCGACCTCCACGGTTT 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 336 ProTrpSerAlaValAla-----LysThrSerProTrpProArg 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 TGCCCT-----CCACGATGTTCTCA----- 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 TrpProGlyMetProAlaGlyProArgProSerCysThrAlaValAlaAlaTrpAlaPro 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 -----GATCTCTGTCAATCATCTTATCAAGG 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 369 AlaArgGlyGlyTyraAlaAlaAlaProAlaArgAsnSerCysSerAlaGlyArgLeuArg 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CATCTGATGTCAAAATTTACTGTCTTCAAGGCTGGGCTTCT-----GCTCTG 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 ArgVal---ArgArgSerAlaArgAlaTrpArgProTrpProAlaGlyArgAlaThr 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 CCGTCAGGACCTTGAGA----- 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 ProAlaSerAspAlaArgArgAspSerAlaSerArgGlyArgProAlaThrAlaAlaAsp 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 -----ACC 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 428 HisProAlaAlaTrpValSerAlaAlaArgArgThrSerSerAlaProIleAlaGlySer 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 TGTCTGGCTCAGAGTACCACTCCGCTGCCGCTGTTCACTGTTCTGTCCAGCA 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 AlaProGlySerGlyThrAlaProArgCysHis-----ProValArgLysArgGlyAla 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 ACTTCATGATCTTCTCACACCGGTGACCTTCCGCTGCGCAGTGC----- 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 466 GlyPro---AlaSerThrGlyArgSerArgArgTrpAlaAlaAlaProAlaArgArg 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 91 CGCCGTCACCCCTATCAGAGAAGCGCTGTGCGGACGTCGCTCGGCGCGGCG 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 485 AlaArgThrGlyIleArgArgThrAlaAlaAlaCysTrpProProProAlaProGly 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 31 -----TCTGCAGTCTTGCTCCGCGCGAAGCA 2
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 505 ArgArgThrProAlaAlaArgArgSerAlaArgThrAlaProArgPro 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

RESULT 15

US-09-252-991A-24055
 / Sequence 24055, Application US/09252991A
 / Patent No. 6551795
 / GENERAL INFORMATION:
 / APPLICANT: Marc J. Rubenfield et al.
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 / FILE REFERENCE: 107196.136
 / CURRENT APPLICATION NUMBER: US/09/252,991A
 / CURRENT FILING DATE: 1999-02-18
 / PRIOR APPLICATION NUMBER: US 60/074,788
 / PRIOR FILING DATE: 1998-02-18
 / PRIOR APPLICATION NUMBER: US 60/094,190
 / PRIOR FILING DATE: 1998-07-27
 / NUMBER OF SEQ ID NOS: 33142
 / SEQ ID NO 24055

/ TYPE: PRT
 / ORGANISM: Pseudomonas aeruginosa
 / US-09-252-991A-24055

Alignment Scores:

Pred. No.:	6.75e-07	Length:	582
Score:	152.50	Matches:	89
Percent Similarity:	30.94%	Conservative:	23
Best Local Similarity:	24.59%	Mismatches:	111
Query Match:	10.07%	Indels:	139
DB:	4	Gaps:	18

US-09-896-522-3 (1-834) x US-09-252-991A-24055 (1-582)

This Page Blank (uspto)

SUMMARIES

8003.075 Million cell updates/sec

Sequence: 1 atggcttcgcgagcgca.....ccagcagcagaccactga 834

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```

-MODEL=rfime+npz.model -DEV=xmln
-O=cgml2/USPTO.spool/US09896522/runat_21112003_184103_2805/app_query.fasta_1.28303
-DB=A_Geneseq_194903 -QFMT=fastan -SUFFIX=+2p.rdg -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bioenum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=5000 -MINLEN=0 -MAXLEN=2000000000
-USERS=US09896522@CG1 1 1.112 @runat_21112003_184103_2805 -ICPU=6
-NO_MAMP -LARGEBINARY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEF TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : A_Geneseq_19Jun03:*

1	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT*
2	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT*
3	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT*
4	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT*
5	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT*
6	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT*
7	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT*
8	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT*
9	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT*
10	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT*
11	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT*
12	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT*
13	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT*
14	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT*
15	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT*
16	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT*
17	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT*
18	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT*
19	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT*
20	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT*
21	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT*
22	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT*
23	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT*
24	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1450	95.8	277	22	AAM39502	Human polypeptide
2	1450	95.8	277	22	AAB93941	Human protein sequ
3	1450	95.8	277	23	AAE16592	Human 57658 prote
4	1450	95.8	296	22	AAM41288	Human polypeptide
5	1372.5	90.7	266	22	AAAG4506	Human uridine kinase
6	1363	90.0	270	22	AAOI4412	Protein of a human
7	1325	87.5	277	22	AAOI4413	protein relating t
8	951	62.8	190	23	ABR93353	Human polypeptide
9	934	61.7	261	22	AAM86394	Human polypeptide
10	934	61.7	261	22	ABP73494	Human transferrase
11	902.5	59.6	337	23	ABP41393	Human ovarian anti
12	897.5	59.3	337	21	AAAB5682	Human prostate car
13	889	58.7	335	22	AAAB4080	Human polypeptide
14	738	48.7	260	22	ABR62307	Drosophila melanog
15	589	38.9	326	22	ABR27215	Novel human diagen
16	579	38.2	120	23	ABG70328	Human MDR protein
17	578	38.2	481	21	AAAG3491	Zea mays protein f
18	578	38.2	490	21	AAAG3490	Zea mays protein f
19	578	38.2	512	21	AAAG34089	Zea mays protein f
20	570.5	37.7	500	21	AAAG47030	Arabidopsis thalia
21	570.5	37.7	511	21	AAAG47029	Arabidopsis thalia
22	545.5	36.0	548	23	AAU98000	Human uridine kinase
23	542	35.8	614	22	ABR60981	Drosophila melanog
24	542	35.8	614	22	ABR67004	Drosophila melanog
25	540.5	35.7	548	23	ABR97433	Novel human prote
26	540.5	35.7	578	22	AAU28181	Novel human secre
27	529	34.9	415	21	AAAG51786	Arabidopsis thalia
28	529	34.9	433	21	AAAG51785	Arabidopsis thalia
29	528	34.9	441	21	AAAG51784	Arabidopsis thalia
30	528	34.9	443	21	AAAG51749	Arabidopsis thalia
31	529	34.9	461	21	AAAG15748	Arabidopsis thalia
32	529	34.9	469	21	AAAG15747	Arabidopsis thalia
33	526.5	34.8	540	21	AAAB41657	Human ORF14221
34	518	34.2	477	21	AAAG45808	Arabidopsis thalia
35	518	34.2	500	21	AAAG45807	Arabidopsis thalia
36	517.5	34.2	466	21	AAAG48809	Arabidopsis thalia
37	507.5	33.5	476	21	AAAG46426	Arabidopsis thalia
38	507.5	33.5	498	21	AAAG30351	Arabidopsis thalia
39	507.5	33.5	499	21	AAAG48425	Arabidopsis thalia
40	507.5	33.5	521	21	AAAG30350	Arabidopsis thalia
41	507	33.5	465	21	AAAG48427	Arabidopsis thalia
42	507	33.5	487	21	AAAG30352	Arabidopsis thalia
43	490.5	32.4	274	21	AAAG52640	Arabidopsis thalia
44	490.5	32.4	292	21	AAAG52639	Arabidopsis thalia
45	490.5	32.4	300	21	AAAG52638	Arabidopsis thalia

RESULT 1

AAM39502

ID AAM39502 standard; Protein; 277 AA.

AC AAM39502;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2647.

XX Human; nocrotropic; immunosuppressant; cyostatic; gene therapy; cancer;

XX peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX leukaemia.

XX Homo sapiens.

OS

ALIGNMENTS

XX
PN NO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Mehran T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; A158658.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2647; 10078bp; English.
XX
CC The invention relates to human nucleic acids (A157798-A161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilization of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 277 AA;
XX
Alignment Scores:
Pred. No.: 1.5e-112 Length: 277
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.77% Indels: 0
DB: 22 Gaps: 0
US-09-896-522-3 (1-834) x AAM39502 (1-277)
QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGGACCGTCCGAC 60
DB 1 MetAlaSerHisIleGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
QY 61 CAGCGGCCCTCTCTGATAGGGGGTGAAGCGCGGCACTGCGACGGGAAGTGAACCTGTGT 120
DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGTTCGTGGGACAGACGAGGTGAACAACGGCGGCGGAGAGTGTGTC 180
DB 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
QY 181 ATCCTGAGCCAGACAGATTCTCAACAAGTCTCTGACGCGACAGACGAGCAAGGACCTTG 240

DB 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGACATGATCAATTTTGAACCATCCAGATGCTTGTGATTAATGATTTGACAGAGACT 300
DB 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 301 CTGAAGACATCGTGGAGGGCAAAACGGTGAAGGTGCCACCTATGATTTTGGACACAC 360
DB 101 LeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrAspPheValThrHis 120
QY 361 TCAAGGTTACCAAGACCAAGGTGTCTACCTGCGAGAGGTGTCTGTGAGGCAATC 420
DB 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
QY 421 TTGGTGTCTACAGCCAGAGAGATCCGGACATGTTCCACTCGGCTCTTCGTGGACAC 480
DB 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCAGCGTCAAGGTCTCTGAAAGAGTTCTCCGGGACGTGCGCGGAGGAGGACCTG 540
DB 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
QY 541 GAGCAGATTCTGACGCGATACACACCTTCGTAAACCGGCTTCGAGGAGTTTGCCTG 600
DB 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGluPheCysLeu 200
QY 601 CCGACAAAGAAGTATCCGATGTGATCATCCACGAGAGGTGACAAATGTGTCGCATC 660
DB 201 ProThrLysLeuTyrLysValIleIleProArgGlyValAspAsnMetValAlaIle 220
QY 661 AACCTGATCGTGCAGACATCCAGACATTCGAAATGTGATCATGCAATGTGACACCGA 720
DB 221 AsnIleIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheAsn 240
QY 721 GGAAGGTCCAAATGGCGGAGCTTACAAAGCGACCTTTTGTAGCCAGGAGCAACCTGG 780
DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGluLysPheHisPro 260
QY 781 ATGCTGACCTTCGGGCAACGGTCACTTTGGAGTCCAGAGCAGACCCAC 831
DB 261 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 277
RESULT 2
AAB93941
ID AAB93941 standard; Protein; 277 AA.
XX
AC AAB93941;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:13952.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

FT Modified-site /label= Casein_kinase_II_phosphorylation_site
 FT 160..163
 FT Modified-site /label= Casein_kinase_II_phosphorylation_site
 FT 167..169
 FT Modified-site /label= Protein_kinase-C_phosphorylation_site
 FT 202..204
 FT Modified-site /label= Protein_kinase-C_phosphorylation_site
 FT 241..246
 FT Modified-site /label= N-myristoylation_site
 FT 247..249
 FT Modified-site /label= Protein_kinase-C_phosphorylation_site
 FT 251..254
 FT Modified-site /label= Casein_kinase_II_phosphorylation_site
 FT 260..265
 FT Modified-site /label= N-myristoylation_site
 FT 264..266
 FT Modified-site /label= Protein_kinase-C_phosphorylation_site
 FT 264..267
 FT Modified-site /label= Amdation_site
 FT 268..271
 FT Modified-site /label= Casein_kinase_II_phosphorylation_site
 FT 273..275
 FT Modified-site /label= Protein_kinase-C_phosphorylation_site
 FT W0200202761-A2.
 PN 10-JAN-2002.
 PD 28-JUN-2001; 2001WO-US21063.
 PP 30-JUN-2000; 2000US-216503P.
 PR (MILL-) MILENNIUM PHARM INC.
 PA
 XX
 XX Gluckmann MA;
 PI WPI; 2002-140091/18.
 XX DR N-PSDB; AAD27186.
 XX PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 XX Claim 4; Fig 1a; 103pp; English.
 PS
 XX The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as hematopoietic neoplastic disorders
 CC (e.g. leukemias), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is human 57658 protein.
 XX
 XX Sequence 277 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.5e-112 Length: 277
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95..77% Indels: 0
 DB: Gaps: 0

US-09-896-522-3 (1-834) x AAB16592 (1-277)
 QY 1 ATGGCTTCGGCGGAGCGGAGACATCCAGAGCCCGCCGAGGCCACCGTCCGAC 60
 Db 1 MetAlaSerAlaGlyGlyGlyAspCysGluSerProAlaProGluAlaAspArgProHis 20
 QY 61 CAGCGGCGCTTCCTGATGAGGGGTGAGCGGCGGACATCGCAGCGGAGATCGACCGTGT 120
 Db 21 GlnArgProPheLeuIleGlyAlaSerGlyGlyThrAlaSerGlyLysSerThrAlaCys 40
 QY 121 GAGAGATCATGAGAGTTGCTGGAGAGAAAGAGTGAACACCGGAGCGGAGAGTGTG 180
 Db 41 GluLysIleIleMetGluLeuLeuGlyGlnAsnGluValGluGluIleArgGlnArgValVal 60
 QY 181 ATCTTACGACGAGACGAGTTCTTACAGGCTCTGACCGGAGACGAGACGAGCTT 240
 Db 61 ILeLeuSerGlnAspArgPheThrLysValIleuThrAlaGluGluLysAlaLysAlaLeu 80
 QY 241 AAGGACAGTACATTTTATGACCATCCAGATGCCCTTGTATATGATTTGACAGACT 300
 Db 81 LysGlyGlnThrAsnPheAspHisProAspAlaPheAspHisPheuMetHisArgThr 100
 QY 301 CTGAGAGACATCGTGGAGGCGAAAGCGTGAAGTCCGACCTATGATTTTGACACAC 360
 Db 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrLysPheValThrHis 120
 QY 361 TCAGGTTACAGAGACACACGCTGCTTACCTCGCGGACGTGGTCTCTTTGAGGCGATC 420
 Db 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuHeGluGlyLe 140
 QY 421 TTGCTGTTCTACAGCGGAGAGATCCGAGACATGTTCCATCGGCTCTTCGTGACACC 480
 Db 141 LeuValPheLysSerGlnIleuIleArgAspMetPheHisIleuArgLeuPheValAspThr 160
 QY 481 GACTCCGACGTACGAGCTGTCTCGAAGAGTTCTCCGAGCGTGGCCGAGAGGACCTG 540
 Db 161 AspSerAspValArgLeuSerArgValLeuArgAspValAlaArgGlyLysArgPheLeu 180
 QY 541 GAGCAGATTCAGACGAGATCCACACCTTCGTGAGCGGCTTCGAGAGATTCGCTCG 600
 Db 181 GluGlnIleuThrIleuThrThrPheValLysProAlaPheGluIleuIlePheCysLeu 200
 QY 601 CCGACAAAGAGATGCCGATGATCATCCACGAGAGAGTGGACAAATGTTGCCATC 660
 Db 201 ProThrLysLysThrAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
 QY 661 AACCTGATCGTGACGACATCCAGACATTCGATGATGTGACATTCGAAATGGACCGA 720
 Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
 QY 721 GAGGGTCCAAATGGGCGGAGCTTCAAGCGGAGACCTTTCTGAGCCAGAGGAGACCCCTGG 780
 Db 241 GlyGlySerAsnGlyArgSerLysArgThrPheSerGluProGlyLysAspProGly 260
 QY 781 ATCTGACCTTCGGAACGATTCATCTGAGTCCAGAGACGAGACCCAC 831
 Db 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerThrProHis 277
 RESULT 4
 ID AAM41288
 AC AAM41288 standard; Protein; 296 AA.
 XX
 XX AAM41288;
 DT 22-OCT-2001 (first entry)
 XX
 XX Human polypeptide SEQ ID NO 6219.
 DE
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KM Leukaemia.
XX Homo sapiens.
OS
XX MO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000MO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR
XX 09-JUL-2000; 2000US-0598042.
PR
XX 19-JUL-2000; 2000US-0620312.
PR
XX 03-AUG-2000; 2000US-0653450.
PR
XX 14-SEP-2000; 2000US-0662191.
PR
XX 19-OCT-2000; 2000US-0693036.
PR
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
DR N-PSDB; AA160444.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6219; 10078bp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA138642-AA142213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
XX
XX
SQ Sequence 296 AA;

Alignment Scores:
Pred. No.: 1.53e-112 Length: 296
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.77% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA141288 (1-296)
QY 1 ATGGCTTCGGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGCGGACCGTCCGCAC 60
DB 20 MetAlaSerIaGlyGluAspCySgIuSerProAlaProGluAlaAspArgProHis 39
QY 61 CAGCGGCGCTTCCTGATAGAGGAGGTAGCGCGGCACTCCGCGGGAAGTGCACCGTGTGT 120
DB 40 GlnArgProheuleuileGlyValSerGlyGlyThrAlaSerGlyIlySerThrValCys 59
QY 121 GAGAGATCATGAGTGTCTGCGGACAGAACGAGGTGGAACGCGCGCAGCGAAAGTGTCTC 180
DB 60 GluIyellleuIetGluIeuIeuGlyGlnAsnGluValGluIuGlnArgGlnArgIyValVal 79

QY 181 ATCTGAGCCAGGACAGGTTCTACAGGTCTTCAGCGGACGAGAGCAAGGCCCTTG 240
DB 80 IleLeuSerGlnAspArgPheIyIyValLeuThrAlaGluIuGlnIySaluIyAlaLeu 99
QY 241 AAAGACAGTACATTTTGGACCATCCAGATGACCTTGTATGATTTGATGACAGACT 300
DB 100 LysGlyGlnIyTrasnPheAspHisProAspAlaPheAspAsnAspIeuMetHisArgThr 119
QY 301 CTGAGAACATCTGGAGGCGCAAAACGGTGAAGTCCGACCTATGATTTGTGACACAC 360
DB 120 LeuIyAsnIleValGluIyLysThrValGluValProThrIyTrasnPheValThrHis 139
QY 361 TCAAGGTACCAAGACCAACGCGTGTCTACCTCCGCGGACGTGTTCTGTTGAGGCATC 420
DB 140 SerArgIeuProGluIyThrValValIyProAlaAspValValIeuPheGluIyIle 159
QY 421 TTGGTGTCTACAGCGGAGATCCGAGACATGTCACCTCGCGCTCTTCGTGAGACACC 480
DB 160 LeuValPheIyTrasnGlnIuIleArgAspMetPheHisIleuArgIeuPheValAspThr 179
QY 481 GACTCCGAGCGTCAGCGTGTCTCGAAGAGTTTCCGGGACGTGCGCCGAGGAGGACCTG 540
DB 180 AspSerAspValArgIeuSerArgValIleuArgAspValArgArgGlyArgAspIeu 199
QY 541 GAGCAGATTCTGACGCGATACACCACTTCGTGAAGCCGCGCTTCGAGAGTCTGCTCG 600
DB 200 GluGlnIleuThrIyTrasnPheValIyProAlaPheGluIyIyPheCysIeu 219
QY 601 CCGACAAAGAGATGCCGATGATCATCCACGAGAGTGGACATATAGTGTGCATC 660
DB 220 ProThrIyLysIyTrasnAspValIleIleProArgGlyValAspAsnMetValAlaIle 239
QY 661 AACCTATCTGTGAGCAGATTCAGACATTTCTGAATGTGACATCTGCAAATGGACCGA 720
DB 240 AsnIeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIySTPHisArg 259
QY 721 GAGGGTCCCAATGGCGGAGCTACAGCGGACCTTTTTCAGGACGAGGACCCCTGGG 780
DB 260 GlyGlySerAsnIyArgSerIyTrasnArgThrPheSerGluProGluIyAspHisProGly 279
QY 781 ATCTGATCTGTGGCAACGATCAGCATTTTGAAGTCCAGACGAGACCCGCAC 831
DB 280 MetIeuThrSerGlyIyAspSerHisIleuGluSerSerArgProHis 296

RESULT 5
AAG64506
ID AAG64506 standard; Protein; 276 AA.
XX
AC AAG64506;
XX
D7 02-OCT-2001 (first entry)
XX
DE Human uridine kinase.
XX
KW Human; uridine kinase; UK.
XX
OS Homo sapiens.
XX
PN CN1287172-A.
XX
PD 14-MAR-2001.
PF 07-SEP-1999; 99CN-0118818.
PR 07-SEP-1999; 99CN-0118818.
XX
PA (UYFU-) UNITV FUDAN.
XX
PI Yu L, Zhao Y, Zhang H;
XX
DR WPI: 2001-409529/44.
DR N-PSDB; AAH5355.

XX Human uridine kinase and its coding sequence, preparation and
PT application -
XX
XX Claim 2, Page 15-16(Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
XX Sequence 276 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 4.39e-106 Length: 276
XX Score: 1372.50 Matches: 266
XX Percent Similarity: 97.11% Conservative: 3
XX Best Local Similarity: 96.03% Mismatches: 7
XX Query Match: 90.65% Indels: 1
XX DB: 22 Gaps: 1
XX
XX US-09-896-522-3 (1-834) x AAG64506 (1-276)
XX
XX 1 ATGCTTCGCGGAGGCGAAGCTGCGAGCCCGCGGAGCCGACGTCGCGAC 60
XX 1 MetAlaSerAlaGlyGlyGlyGlyAlaArgAlaArgAla--GlyAlaAsnArgProHis 19
XX
XX 61 CAGCGCCCTTCCTGATAGGGGTAGCGCGCACTGCCAGCGGAAGTCGCGTGT 120
XX 20 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 39
XX
XX 121 GAGAGATCATGAGTTCCTGGACAGAACAGGTGGAAACGCGGCGGAGGTGTC 180
XX 40 GluLysIleMetGluLeuGlyGlnAsnGlyValGlnGlnArgHisGlyLysValVal 59
XX
XX 181 ATCTGAGCCGAGCAGGCTTCAACAGGTCTGACGCGAGAGCAGAACGCGACGCTTG 240
XX 60 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 79
XX
XX 241 AAAGAGACGTAACAATTTGACCATCCAGATGCTTTGATAATGATTGATGACAGACT 300
XX 80 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 99
XX
XX 301 CTGAGAGAACATCGTGGAGGGGAAAACGCTGAGGTGCCGACCTTGTGATTTGTACAC 360
XX 100 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 119
XX
XX 361 TCAAGTTACCGAGACGACGCGTGTCTACCTCGGACGCGTCTGTTGAGGCGATC 420
XX 120 SerArgLeuProGlnThrValValTyrProAlaAspValValLeuPheGlnGlyIle 139
XX
XX 421 TTGGTGTCTACACCGACGAGATCCGGACATGTTCACCTGCGCTTCTGAGACAC 480
XX 140 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisIleuArgLeuPheValAspThr 159
XX
XX 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGCGCGCCGAGGAGGACTCG 540
XX 160 AspSerAspValAlaGlnSerArgArgValLeuArgAspValAlaArgGlyAlaArgAspLeu 179
XX
XX 541 GAGCAGATTGACGACGATACACACCTTCGTAAGCGCGGCTTCGAGAGTTTGCCTG 600
XX 180 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 199
XX
XX 601 CCGACAAAGAGATGCCGATGTATCATCCACGAGAGTGAACAATATGTTGCCATC 660
XX 200 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAsnMetValAlaIle 219
XX
XX 661 AACCTGATCGGACGACATCCAGACATTTGATGATGTCATCTGCAATGGCACCGA 720
XX 220 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnAspIleCysLysTyrPheArg 239
XX
XX 721 GGAAGGTCCAAATGGGCGGAGTACAAAGCGACCTTTCTGAGCGACGAGGGACCACTGGG 780
XX 240 GlyLysLeuArgHisGlyArgHisTyrLysArgThrPheSerGlnProGlyAspHisProGly 259
XX
XX 781 ATGCTGACCTCTGGCAAAACGCTCAATTGGAGTCCAGACAGACAGACCCAC 831

DB 260 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 276
|||||
RESULT 6
ID AA014412
XX AA014412 standard; Protein; 260 AA.
XX
XX AA014412;
XX
XX 02-MAY-2002 (first entry)
XX
XX DE Protein of a human uridine kinase (UDK).
XX
XX DE Human; uridine kinase; diagnostic assay; mutation detection; UDK;
XX KW probe; chromosome localization study; tissue expression; gene therapy;
XX KW antibody; vaccine; human ovarian cancer; immunological disorder;
XX KW human colon carcinoma; immunogen.
XX
XX OS Homo sapiens.
XX
XX OS
XX PN W0200172963-A2.
XX
XX PD 04-OCT-2001.
XX
XX PF 27-MAR-2001; 2001WO-US09663.
XX
XX PR 27-MAR-2000; 2000US-0536647.
XX
XX PA (SMIRK) SMITHKLINE BEECHAM CORP.
XX
XX PI Ho YS, Johnson RK;
XX
XX DR WPI: 2001-626259/72.
XX
XX DR N-PDB: AAK98735.
XX
XX PT Novel human uridine kinase polypeptides useful for treating cancers,
XX and to identify agonists and antagonists of the polypeptide useful for
XX PT treating conditions associated with uridine kinase imbalance -
XX
XX PS Claim 3; Page 29-30; 31pp; English.
XX
XX CC The invention relates to newly identified human uridine kinase (UDK)
XX CC polypeptides and polynucleotides and methods for producing such
XX CC polypeptides by recombinant techniques. Also disclosed in the invention
XX CC are methods for utilizing uridine kinase polypeptides and polynucleotides
XX CC in diagnostic assays. The polynucleotides and polypeptides of the
XX CC invention may be used as diagnostic reagents by detecting mutations in an
XX CC associated gene. An array of oligonucleotide probes comprising the
XX CC uridine kinase polynucleotide sequence or fragments thereof can be
XX CC constructed to conduct efficient screening of genetic mutations, for
XX CC example. Detection of abnormally decreased or increased levels of
XX CC polypeptide or mRNA expression may also be used for diagnosing or
XX CC determining susceptibility of a subject to a disease of the invention.
XX CC The polynucleotide sequences of the invention can be used for chromosome
XX CC localisation studies and tissue expression studies. The polypeptides of
XX CC the invention or fragments thereof may be used as immunogens to produce
XX CC antibodies. These antibodies may be employed to isolate or identify
XX CC clones expressing the polypeptide. The polypeptides and polynucleotides
XX CC of the invention can be used as a vaccine or in gene therapy to treat
XX CC diseases such as human ovarian cancer, human colon carcinomas, and
XX CC immunological disorders. This sequence represents the protein of a human
XX CC uridine kinase of the invention.
XX
XX SQ Sequence 260 AA;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2.67e-105 Length: 260
XX Score: 1363.00 Matches: 260
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 90.03% Indels: 0
XX DB: 22 Gaps: 0

US-09-896-522-3 (1-834) x AA014412 (1-260)

```

QY 1 ATGGCTTCGGCGGAGGCGAAGCTGGAGAGCCCGCGCGAGCGCCGACCGTCCGAC 60
DB 1 MetAlaSerIaGlyGluAspCysGluSerProIaProGluIaAspArgProHis 20
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGCACTCCAGCGGGAAGTGAACCGTGTG 120
DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrIaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGTG 180
DB 41 GlnLysIleMetGluLeuLeuGlnIaAsnGlnValGluGlnArgGlnArgValVal 60
QY 181 ATCTGAGCGGAGAGCAGGTTCTCAAGGTCCTGACGCGGACAGAGGAGGCGGCTTG 240
DB 61 IleuSerGlnAspArgPheTyrLysValLeuThrIaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGCAGTACATTTTGAACCATCCAGATCCCTTTGATGATGATGATGACAGGACT 300
DB 81 LysGlyGlnIleTyrAsnProPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 301 CTGAGAGAACCTGCTGAGGGGCAAAACGCTGAGGTCGCACTTATGATTTTGTGACAC 360
DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
QY 361 TCAAGGTTACACAGACACGAGTGTCTACCTCCGAGCGTGTCTGTTGAGGGGATC 420
DB 121 SerArgLeuProGluThrIaValIaValTyrProIaAspValIaLeuPheGluGlyLe 140
QY 421 TTGGGTGTTACAGCCAGGAGATCCGGGACATGTTCCAGCTGGCGCTCTTGAGGACAC 480
DB 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 481 GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTCGCGCGAGGAGGACCTG 540
DB 161 AspSerAspValIaArgLeuSerAspArgValIaLeuArgAspValIaArgGlyIaArgAspLeu 180
QY 541 GACCAAGTTCTGACGCGATACACCACTTGTGTAAGCGCGCTTTCGAGGAGTCTGCTG 600
DB 181 GlnGlnIleLeuThrGlnTyrIaThrPheValLysProIaPheGluGlnPheCysLeu 200
QY 601 CCGACAAAGAGATGATGCGGATGATCATCCACGAGAGATGAGCAATATGTTGCCATC 660
DB 201 ProThrLysLysTyrIaAspValIleIleProArgGlyValAspMetValIaIle 220
QY 661 AACCTGATCGTGACGACATCCAGGACATTTCTGATGTCGATCTGCAATGGACCGA 720
DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheArg 240
QY 721 GAGGGTCCCAATGGCGGAGACTACAGCGGACCTTTCTGAGCCAGGGGACACCGCTGG 780
DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260

```

RESULT 7

ID AA014413 standard; Protein; 277 AA.

AC AA014413;

DT 02-MAY-2002 (first entry)

DE Protein relating to a human uridine kinase (UDK) of the invention.

XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;

KW probe; chromosome localisation study; tissue expression; gene therapy;

KM antibody; vaccine; human ovarian cancer; immunological disorder;

XX human colon carcinoma; immunogen.

OS Unidentified.

XX WO20012963-A2.

```

PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001MO-US09663.
XX
PR 27-MAR-2000; 2000US-0536647.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Ho YS, Johnson RK,
XX
DR WPI; 2001-626259/72.
XX
PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
XX
PS Disclosure; Page 23; 31pp; English.
XX
CC The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC NOTE: The present sequence is stated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.
XX
SQ Sequence 277 AA:

```

Alignment Scores:

Pred. No.:	4,04e-102	Length:	277
Score:	1325.00	Matches:	254
Percent Similarity:	95.67%	Conservative:	11
Best Local Similarity:	91.70%	Mismatches:	12
Query Match:	87.52%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-3 (1-834) x AA014413 (1-277)

```

QY 1 ATGGCTTCGGCGGAGGCGAAGCTGGAGAGCCCGCGCGAGCGCCGACCGTCCGAC 60
DB 1 MetAlaSerIaGlyGlyGlySerGluSerIaIaIaProGluIaAspArgProGln 20
QY 61 CAGCGGCCCTTCCTGATAGGGGTGAGCGGCGGCACTCCAGCGGGAAGTGAACCGTGTG 120
DB 21 ProArgProPheLeuIleGlyValSerGlyGlyThrIaSerGlyLysSerThrValCys 40
QY 121 GAGAGATCATGAGATTGCTGGGACAGAACGAGGTGGAACGCGGACCGGAAGTGTG 180
DB 41 GlnLysIleMetGluLeuLeuGlnIaAsnGlnValIaAspArgGlnArgValVal 60
QY 181 ATCTGAGCGGAGAGCAGGTTCTCAAGGTCCTGACGCGGACAGAGGAGGCGGCTTG 240
DB 61 IleuSerGlnAspArgPheTyrLysValLeuThrIaGluGlnLysAlaLysAlaLeu 80
QY 241 AAAGCAGTACATTTTGAACCATCCAGATCCCTTTGATGATGATGATGACAGGACT 300

```

Db LysGlyGlnItyrAenPheAspHisProAspAlaPheAspAenAspLeuMetHisIleYThr 100
Qy 301 CTGAGAGAACCTGTCGAGGGGAAACGTTGGAGTCCGACCTTGAATTTTGTACACAC 360
Db 101 LeuYsAsnIleValIGluGlyLysThrValIGluValProThrItyrAspPheValThrHis 120
Qy 361 TCAAGGTTACAGAGACACCGGTGTCTACCTCCGACGTCGTCTGTTGAGGAGCATC 420
Db 121 SerArgLeuProGluThrThrValValItyrProAlaAspValIleLeuPheGluGlyIle 140
Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTGACACC 480
Db 141 LeuValPheItyrThrGlnGluIleAlaArgAspMetPheHisIleuArgLeuPheValAspThr 160
Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGGACGCGCCCGACGAGAGGACCTG 540
Db 161 AspSerAspValAlaArgLeuSerArgArgValIleuArgAspValIGlnArgGlyArgAspLeu 180
Qy 541 GAGCAGATTCTGACGACATACACACCTTCCTGAAGCCGCGCTTCGAGAGTTTGCCTG 600
Db 181 GluGlnIleLeuThrGlnItyrThrAlaPheValItyrProAlaPheGluGluPheCysLeu 200
Qy 601 CCGACAAAGAAGATGCCGATGTATCATCCACGAGAGTGAACAATGCTTGCATC 660
Db 201 ProThrItyrLysItyrAlaAspValIleIleProArgGlyValAspAspMetValAlaIle 220
Qy 661 AACCTGATCGTCGACATCCAGACATCTGATGATGTCATCTGCAATGGCACCGA 720
Db 221 AsnIleuIleValIGlnHisIleGlnAspIleLeuAsnGlyAspLeuCysLysArgHisArg 240
Qy 721 GGAGGCTCCAAATGGCGGAGCTACAGCGGACCTTTCTGAGCGAGGAGCACCTGGG 780
Db 241 GlyGlyProAsnGlyArgAsnHisIleYThrGlnPheProGluProGlyAspHisProGly 260
Qy 781 ATGCTGACCTTGGCAACCGTCACTTGGATGATGTCAGCAGACACCCAC 831
Db 261 ValIleuAlaThrGlyLysArgSerHisIleuGlnIleuSerSerArgProHis 277
RESULT 8
ABB89353
ID ABB89353 standard; Protein; 190 AA.
AC ABB89353;
XX 24-MAY-2002 (first entry)
DT Human polypeptide SEQ ID NO 1729.
DE
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antiallergic; hepatocrotropic; antidiabetic; antiinflammatory; antitumor;
KM valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KM neurological disease; infection; human; secreted protein.
XX
XX Homo sapiens.
OS
XX MO2001.90304-A2.
PN
XX 29-NOV-2001.
PD
XX 18-MAY-2001; 2001WO-US16450.
PF
XX 19-MAY-2000; 2000US-205515P.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Birse CE, Rosen CA;
PI
XX WPI; 2002-122018/16.
DR N-PSDB; ABL89762.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,

PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
PS Claim 11, SEQ ID NO 1729; 2081bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL9449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 190 AA:
Alignment Scores:
Pred. No.: 5.73e-71 Length: 190
Score: 951.00 Matches: 187
Percent Similarity: 97.91% Conservative: 0
Best Local Similarity: 97.91% Mismatches: 3
Query Match: 62.81% Indels: 1
DB: Gaps: 0
US-09-896-522-3 (1-834) x ABB89353 (1-190).
Qy 1 ATGCGTTCGCGGAGGCGAAGATCGAGAGCCCGCGGAGCGACCTCCGAC 60
Db 1 MetAlaSerIleGlyGluAspCysGluSerProAlaProIleuAlaAspArgProHis 20
Qy 61 CACGGGCCCTTCTGATPAGGGGTGAGCGGCGGACCTGCCAGCGGAAGTGCACCTGT 120
Db 21 GlnArgProPheLeuIleGlyValSerGlyItyr**SerGlySerThrValCys 40
Qy 121 GAGAAGATCATGGAGTTCGTCGAGACGAGACGAGTGAACGCGGAGCGGAAGTGGTC 180
Db 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValIGluGlnArgIleValIle 60
Qy 181 ATCTGAGCCAGAGACAGTTCCTCAAGATCTTGACCGCAGAGCAGAGCCAGCCTTG 240
Db 61 IleuSerGlnAspArgPheItyrLysValIleuThrAlaGluGlnItyrAlaLysAlaLeu 80
Qy 241 AAGAGACATGATCAATTTTGAATCCATGCTGCTTGAATGATTTGATGCACAGACT 300
Db 81 LysGlyGlnItyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Qy 301 CTGAGAGAACCTGTCGAGGGGAAACGTTGGAGTCCGACCTTGAATTTTGTACACAC 360
Db 101 LeuYsAsnIleValIGluGlyLysThrValIGluValProThrItyrAspPheValThrHis 120
Qy 361 TCAAGGTTACAGAGACACCGGTGTCTACCTCCGACGTCGTCTGTTGAGGAGCATC 420
Db 121 SerArgLeuProGluThrThrValValItyrProAlaAspValIleLeuPheGluGlyIle 140
Qy 421 TTGGTGTTCACAGCCAGAGATCCGGACATGTTTCCACTGCGCTCTTCTGTGACACC 480
Db 141 LeuValPheItyrSerGlnGluIleAlaArgAspMetPheHisIleuArgLeuPheValAspThr 160
Qy 481 GACTCCGAGCTGAGGCTGTCTCGAAGAGTTCTCCGGACGCGCCCGACGAGAGGACCTG 540
Db 161 AspSerAspValAlaArgLeuSerArgArgValIleuArgAspValIGlnArgGlyArgAspLeu 180
Qy 541 GACGAGATTCTGACGACATACACACCTTCCTGAAGCCGCGCTTCGAGAGTTTGCCTG 600

Db 181 GluGlnIleuThrhHis-ThrProProSer 190

RESULT 9

ID AAM38694 standard; Protein; 261 AA.

XX AAM38694;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 1839.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia.

XX Homo sapiens.

XX WO20015312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 XX Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX MPI; 2001-442253/47.

XX N-PSDB; AAI57850.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -

XX Example 3; SEQ ID NO 1839; 10078bp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
 XX the encoded polypeptides (AAM38642-AA42213) with nootropic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localised neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: Immune system suppression.
 XX Activation/inhibition of chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukaemia and
 XX C.N.S disorders.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification.

XX Sequence 261 AA;

Alignment Scores: 1.66e-69 Length: 261
 Pred. No.: 934.00 Matches: 189
 Score:

Percent Similarity: 79.41% Conservative: 27
 Best Local Similarity: 69.49% Mismatches: 40
 Query Match: 61.69% Indels: 16
 DB: 22 Gaps: 4

US-09-896-522-3 (1-834) x AAM38694 (1-261)

QY 22 GACTCGAGAGCCCGCGCGGAGCCGACCGTCCGAC---CAGCGCCCTTCTGATTA 78
 DB 4 AepSerGIuGlnThrIeuGlnAenHISgInGInProAsnGIyGIuGluProPheLeuIle 23
 QY GGGGTGAGCGCGCGGACCTCCAGCGGAAAGTGAACCGTGTGAGAAGATCATGAGATTG 138
 DB 24 GIyValSerGIyGIyThrAlaSerGIyLysSerSerValCyaAlaLysIleValGInLeu 43
 QY 139 CTGGGACAGAGAGGTGGAACCGCGGACCGGAGAGTGTGATCTTGAGCCAGGACAG 198
 DB 44 LeuGIyGlnAenGlnValAspTyrArgGInLysGlnValIleLeuSerGlnAepSer 63
 QY 199 TTCTACAAAGTCTCGACGCGGAGGAGGCGGAGGCGCTTGAAGAGACATGATTT 258
 DB 64 PheTyrArgValLeuThrSerGIuGlnLysAlaLysAlaLeuLysGIyGlnPheAsnPe 83
 QY 259 GACCATCCAGATGCTTTGATATGATTTGATGACAGGACTCTGAGAACAATCGTGAG 318
 DB 84 AspHisProAspAlaPheAspAsnGInLysIleLeuLysThrLeuLysGIuIleThrGIn 103
 QY 319 GGCAGAAACGTTGAGGTGCGGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACC 378
 DB 104 GIyLysThrValGInIleProValTyrAspPheValSerHisSerArgLysGIuGInThr 123
 QY 379 ACGGTGCTTACCGTCCGAGCGGTGTTGTTGAGGCGCATGTTGTTTCAACGCCAG 438
 DB 124 ValThrValTyrProAlaAspValValLeuPheGInLysIleLeuAlaPheTyrSerGIn 143
 QY 439 GAGATCCGAGACATGTTCCACCTGCTTCTGTTGAGACCGGACTCCGACCTGAGGCTG 498
 DB 144 GIuValArgAspLeuPheGlnIleMetLysLeuPheValAspThrAspAlaSerHisArgLeu 163
 QY 499 TTCGAGAGATTCTCCGGAGCGT---CGCCAGAGGAGGACCTGAGGACATTTGACG 555
 DB 164 SerArgArgValLeuArgAspIleSerGIuArgGIyArgAspLeuGInIleLeuSer 183
 QY 556 CAGTACACACCTGCTGAGCGCGGCTTCCAGGAGTCTGCTGCCAAGAAAGTAT 615
 DB 184 GlnTyrIleThrPheValLysProAlaPheGInGInPheCysLeuProThrLysLysTyr 203
 QY 616 GCCGATGATCATCCAGAGAGTGAATATGTTGTCATCACTGATCGGCGAG 675
 DB 204 AlaAspValIleIleProArgGIyAlaAspAsnLeuValAlaIleAsnLeuIleValGIn 223
 QY 676 CACATCCAGACATTTCTGAATGTTGATCTGCAATGCGACCGAGAGGTCATATGG 735
 DB 224 HisIleGlnAspIleLeuAsnGIy-----Gly 232
 QY 736 CCGAGTTACAACGACCTTTTCTGAGCGAGGAGCAACCTGGAGTGTGACTTGTGC 795
 DB 233 ProSerLysArgGlnThr-----AsnGIyCysLeuAsnGIyTyrThrProSerArg 249
 QY 796 AAACGTCATTTGGAGTCCAGGACAGACCCAC 831
 DB 250 LysAlaGlnAlaSerGIuSerSerArgProHis 261
 RESULT 10
 AAB73494
 ID AAB73494 standard; Protein; 261 AA.
 XX AAB73494;
 XX 31-JUL-2001 (first entry)
 DE Human transferase HTFS-1, SEQ ID NO:1.
 XX

XX	Human transference; HTFS; agonist; antagonist; cellular signalling;
KM	proliferation; cell proliferative disorder; immune disorder;
KM	atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM	inflammation; AIDS; Addison's disease; allergy; ascuma; anaemia;
KM	cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM	multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM	systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM	hemodialysis; extracorporeal circulation; trauma; transgenic animal;
gene	therapy; drug screening.
OS	Homo sapiens.
XX	
XX	WO200132888-A2.
XX	
XX	10-MAY-2001.
PD	
XX	
PF	02-NOV-2000; 2000WO-US0485.
XX	
XX	04-NOV-1999; 99US-0163595.
PA	
XX	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Tang YF, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI	Smith LL, Azimzai Y, Lu DM, Baughn MR;
XX	
DR	WPI, 2001-328796/34.
XX	
XX	N-PSDB; AAK33801.
PT	Human transference polypeptides and polynucleotides useful for
PT	diagnosis, prevention and treatment of cell proliferative and immune
PT	system disorders and for identifying agonists and antagonists -
DS	Claim 1; Page 103-104; 157pp; English.

CC Sequencec AAB75342-AAB75355 represent novel human transferase proteins
CC HTF5-1 to HTF5-42 and sequencec AAB23801-AAB23842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HTFS proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferase activity, or for compounds that specifically bind to an HTFS
CC protein or which modulates the activity of an HTFS protein.
CC Pharmaceutical compositions comprising an HTFS protein, HTFS
CC agonist or antagonist, or genetic construct encoding an HTFS
CC protein are useful for treating a disease or condition associated
CC with decreased or increased such expression of functional HTFS. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation, CC
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HTFS may be
CC used for the diagnosis of disorders associated with the expression of
CC HTFS, or in assays to monitor patients being treated with HTFS or
CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.

SQ Sequence 261 AA;

Alignment Scores:

Pred. No.: 1.66e-69

Score:

Percent Similarity: 79.41%

Best Local Similarity: 69.49%

Query Match:	61.69%	Indels:	16
DB:	22	Gaps:	4

US-09-896-522-3 (1-834) x AAB73494 (1-261)

22 GACTGCGAGAGCCCCGCGCGGAGAGCCGACCGTCCGAC --- CAGCGCCCTTCTGATA 78

Db 4 AspSerGluGlnThrLeuGlnAsnHisGlnGlnProAsnGlyGlyGluProPheLeuIle 23

79 GGGGTGAGCGCGCACTGCCAGCGGAAGTCGACCGTGTGTGAGAAGATCATGGAGTTG 138

Db 24 GlyValSerGlyGlyThrAlaSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43

139 CTGGACAGACGAGGTGGAACAGCGGACGGAGGTGTCTCTGAGCCAGGACAGG 198

Db 44 LeuGlyGlnAsnGluValAspTyrArgGlnIlyscIlnValValIleLeuSerGlnAspSer 63

199 TTCTACAAGTCTCTGACGGCAGAGCAGAAGGCCAAGGCTTGAAAGGACAGTACAATTT 258

259 GACCATCCAGATGCCCTTGATAATGATTGATGCACAGGACTCTGAAGACATCGTGAG 316

Db 84 AspHisProAspAlaPheAspAsnGluLeuIleLeuLysThrLeuLysGluIleThrGlu 103

319 GGCAAAACGGTGGAGGTGCCGACCTATGATTTTGTGACACACTCAAGGTTACCGAGACC 378

Db 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGluGluThr 123

379 ACGGTGCTACCCCTGCCGACGTGTTCTGTTGAGGGCATCTTGGTCTACAGCCAG 438

Db 124 ValThrValTyrProAlaAspValValLeuPheGluGlyIleLeuAlaPheTyrSerGln 143

439 GAGATCCGGACATGTTCCACCTGCGCCTCTTCGTGGACACCGACTCCGACGTCAGGCTG 498

Db 144 GluValArgAspLeuPheGlnMetLysLeuPheValAspThrArgLeu 163

499 TCTCGAAGAGTTCTCCGGGACGTG--CGCCGAGGGAGGACCTGGAGCAGATTCTGACG 555

Db 164 SerArgArgValLeuArgAspIleSerGluArgGlyArgAspLeuGlnIleLeuSer 183

556 CAGTACACCACCTTCGTGAAGCCGGCCTTCGAGGAGTTCTGCCTGCCGACAAGAAGTAT 611

Db 184 GlnTyrIleThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLysTyr 203

616 GCCGATGTGATCATCCACGAGGAGTGCACAATATGTTGCCATCAACCTGATCGTGCAG 675

Db 204 Alaaspvalllelleproargglyalaaspasnlleuvalalaileasnleuilevalgin 223

676 CACATCCAGGACATTCTGAATGGTGACATCTGCAATGGCACCAGGAGGTCCAATGGC 735

Db 224 HisIleGlnAspIleLeuAsnGly-----Gly 232

736 CGAGCTACAGCGACCTTTCTGAGCCAGGGGACCCTGGATGCTGACCTCTGGC 7955

Db 233 ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrThrProSerArg 249

QY 796 AAACGGTCACATTTGGAGTCCAGCAGCAGACCCAC 831

Db 250 LysArgGlnAlaSerGluSerSerArgProHis 261

RESULT 11
ABD41303

ID	ABP41393 standard; Protein; 337 AA.
XX	

AC ABP41393;
XY

DT 22-AUG-2002 (first entry)

Human ovarian antigen HOELP29, SEQ ID NO:2525.

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW

XX interfertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW gastrointestinal disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 1p32.
XX
XX Homo sapiens.
OS
XX
XX WO200200677-A1.
XX
XX
XX 03-JAN-2002.
XX
XX
XX 07-JUN-2001; 2001WO-US18569.
XX
XX 07-JUN-2000; 2000US-209467P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Biase CE, Rosen CA,
PI
XX
XX WPI: 2002-147878/19.
XX
XX N-PSDB; AB054470.
XX
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
XX
XX Claim 11; SEQ ID No 2525; 2922pp; English.

Alignment Scores:	
Pred. No.:	7,62e-67
Score:	907.50
Percent Similarity:	73.65%
Best Local Similarity:	64.86%
Query Match:	59,61%
DB:	23
Length:	337
Matches:	102
Conservative:	26
Mismatches:	44
Indels:	3
Gaps:	6

US-09-896-522-3 (1-834) X ABP41393 (1-337)

[illegible]

	RESULT 12
AAB56582	
ID AAB56582 standard; Protein; 337 AA.	
XX	
AC	
AAB56582;	-
XX	
DT 13-MAR-2001 (first entry)	
XX	
DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.	
XX	
TX Human; prostate cancer; prostate cancer antigen; detection; diagnosis	

PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
XX	
PA	(HYSEQ-) HYSEQ INC.
XX	
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI	Wang U, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI; 2001-442253/47.
DR	N-PsDB; AA159636.
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
XX	
PS	Example 2; SEQ ID NO 5411; 10078pp; English.
CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AA138642-AA442213) with nucleotide,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nerve injuries, peripheral neuropathy and
CC	localized neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilization of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemia and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SQ	Sequence 335 AA.
Alignment Scores:	
Pred. No.:	1.02e-65 Length: 335
Score:	889.00 Matches: 190
Percent Similarity:	72.48% Conservative: 26
Best Local Similarity:	63.76% Mismatches: 46
Query Match:	58.72% Indels: 36
DB:	Gaps: 7
US-09-896-522-3 (1-834) x AM40480 (1-335)	
OY	4 GCTTCGGCGGAGCGCA----- 21
Dd	52 AlacIylalagIlglylnarGalaserValArgThrgIyserglYnrgArGlYgIYArg 71
OY	22 -----GATCGCAGACGCCCGCGCGAGCCGACCGTCCGCAC---CACCG 66
Dd	72 ThrMetAlagIYApsSerIugInthrLeuGlnAsnHisGlnIndIpProAnsgIYgIYgln 91
OY	67 CCTTCCTCATGTAGGGGTGAGCGCGCGACACTGCACAGCGGAATGCACCGGTGTGAAGAAG 126
Dd	92 ProPheLeuIlleISclYAlserIgtYgnrAlaserGIlysserServalYcSalAllys 111
OY	127 ATCATGAGACTTGCTGGAGCAGAAcGAGGTGAAcACGCGcGACGGAAGGTGATCTCTG 166
Dd	112 IlleValGlnLeuEngIYgInAnsgIlnValApyTYArGInIlyngIlnValIleleu 131
OY	187 AGCCAGACAGGTTCTACAAGCTCTTCAGCGCAGACACAGAAggCCAAggCTTGAAAGCA 246
Dd	132 SerGlnApsSerPheYnrgValleutnrSetGInIlynsAlaySalaleuLYsgIY 151
OY	247 CAGTACAAATTGTTGACCATCCAGATGCGCTTGTGAATAATTTGATGACAGAGCTTGAAG 306
Dd	152 GlnPheAnsrPheAsphIsrProAspAlaPheNdePasnGlnIleuIleuYstrIrleuLys 171
OY	307 AACATCGTGGAGGCGCAAAcGCTGGAGGTGCGCACTATGATTTTGTGACACCTCAAGG 366

[illegible]

capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-ABB72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 260 AA:

Alignment Scores:

Pred. No.:	3,73e-53	Length:	260
Score:	738.00	Matches:	143
Percent Similarity:	82.94%	Conservative:	32
Best Local Similarity:	67.77%	Mismatches:	34
Query Match:	48.75%	Indels:	2
DB:	22	Gaps:	2

US-09-896-522-3 (1-834) x ABB62307 (1-260)

QY 67 CCTTCTGATAGGAGGAGCGCGGCACTGCCAGCGGAAGTCCAGCCGTGTGAGAAG 126
 |||||
 Db 27 Prothetleuileglyvalalaglyglythralaserglyserthrvalcyselyls 46
 |||||
 QY 127 ATCATGAGTTCGCGGAGACAGAGAGTGAACAGCGCGCGGAAGTGTATCCTG 186
 |||||
 Db 47 Ilelecgulnleuglylmalaglwetaspshthrglnaaglnvalvalserlle 66
 |||||
 QY 187 AGCCAGGAGAGTTCACAGAGTCTGACGAGCAGAGCAAGGCCAGCTTGAAGA 246
 |||||
 Db 67 Serlgnapsrserphetyrarglulethrproalaglwulysalalyalaglnlygly 86
 |||||
 QY 247 CAGTACATTTTGAACATCCAGATGCTTGTATATGATTTGATGACAGAGCTTGAAG 306
 |||||
 Db 87 Leuheaenpheaephlaeproaspalapheasnlgulnleumetyrserthrleugln 106
 |||||
 QY 307 AACCTCGAGGAGGCAAAAGCGTGGAGGCGGACCTATGATTTTGACACACTCA--- 363
 |||||
 Db 107 Asnilleuulysglyhslsvalglnleuproserlyrasyrlyrargthrasnserleu 126
 |||||
 QY 364 AGGTACAGAGACACAGCGTGTACCTCGAGAGCGTGTCTTGTGAGGACCTTG 423
 |||||
 Db 127 Asphegluamvalleuvalilletyrprohlaaspralvalleuphegluglylleleu 146
 |||||
 QY 424 GTGTTTACAGCGAGAGATCCGGAACATGTTCCACCTGCGCTTCTGTGACACCGAC 483
 |||||
 Db 147 Valpheytyrphetpolsylearglileuphehismetlyslleuphevalasprthr 166
 |||||
 QY 484 TCCGACGTCGCGTCTCGAAGAGTTCTCCGAGACGCG---CGCCAGGAGGAGACTG 540
 |||||
 Db 167 Seraspthrphalaleuvalargvalproargpaleuaglnuaglyargaspheu 186
 |||||
 QY 541 GAGCAGATTCTGACGACATACACACCTTCGTGAAGCGGCTTCGAGAGTTGTGCTG 600
 |||||
 Db 187 Aspalvalleuthrnglnlymerthrphetalysprohlaaspraleugluuphecyser 206
 |||||
 QY 601 CCGCAAAAGATGCCGATGTATCATCCACGAGAGTGGACAATATGTTGCCATC 660
 |||||
 Db 207 Prothrlyslsyphealaaspyalleleleproargglyalaaspaantrvalalaile 226
 |||||
 QY 661 AACCTGATCGTGCAGCATCCAGACATCTTG 693
 |||||
 Db 227 Aspneuilevalhshstilleglyluleu 237
 |||||

RESULT 15
 ABG27216 standard; Protein, 326 AA.
 ID ABG27216
 AC ABG27216;
 XX

DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27207.
 XX
 KW Human, chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dermanac RT, Liu C, Tang YT;
 XX
 DR WPI, 2001-639362/73.
 DR N-PSDB; AAS91403.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 20; SEQ ID No 57575; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 326 AA:

Alignment Scores:

Pred. No.:	1.08e-40	Length:	326
Score:	589.00 <td>Matches:</td> <td>132</td>	Matches:	132
Percent Similarity:	63.64% <td>Conservative:</td> <td>29</td>	Conservative:	29
Best Local Similarity:	52.17% <td>Mismatches:</td> <td>52</td>	Mismatches:	52
Query Match:	38.90% <td>Indels:</td> <td>40</td>	Indels:	40
DB:	22	Gaps:	7

US-09-896-522-3 (1-834) x ABG27216 (1-326)

QY 4 GCTTCGCGGAGGCGAGAGTTCGAGAGCCCGCGCG---GAGCCGACCGTCCGCAC 60
 |||||
 Db 67 Alaglnalalaglyglylvalalagluuprotprpoglythrhlaserhrprocy 86
 |||||
 QY 61 CAG-----CGGCCCTTCTGATAGGAGTGAACGCGGCGACT 96
 |||||
 Db 87 ArgThrThrsererprothrhlalaserthrserleuileglyalthrtrpglythr 106
 |||||

```
Qy 97 GCCAGCGGGAAGTCGACCGTGTGTGAGAGATCATGGAGTTGCTGGGACAGAAAGAGGTG 156
Db 107 AlasercilysserservalCysAlaLysIleValGlnLeuLeuGlyGlnAsnGluVal 126
Qy 157 GAAACAGCGGAGCGAGAGTGTGTCATCTGAGCCAGGACAGATTCTACAAGTCTTGACG 216
Db 127 AspTyrArgGlnLysGlnValValIleLeuSerGlnAspSerPheTyrArgValLeuThr 146
Qy 217 GCAGAGCAGAAAGCGCAAGCGCTTGAAAGACAGTACAAATTTTGACCATCCAGATCCCTT 276
Db 147 SerGlnGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPheAspHisProAspAlaPhe 166
Qy 277 GATATGATTGTTGACACAGACCTGTAAGAACATCGTGAGGCGCAAAACGTGAGAGTG 336
Db 167 AspAsnGlnLeuLysIleLeuLysThrLeuLysGlnIleThrGlnGlyLysThrValGlnIle 186
Qy 337 CCGACCTATGATTTTGTGACACACTCAAGGTTACAGAGACCAAGGTGTCTACCTGCG 396
Db 187 ProValTyrAspPheValSerHisSerArgLysGlnGluThrValThrValTyrProAla 206
Qy 397 GACGTGTTCTGTTTGAGGGCATCTTGTTCTTACAGCCAGAGATCCGGACATGTTG 456
Db 207 AspValValLeuPheGlnGlyIleLeuAlaPheTyrSerGlnGlnValArgAspLeuPhe 226
Qy 457 CACCTGGCCTCTTGTGGAACCGACTCCGAGCTCAGGCTGTCTCGAGA---GTTCTC 513
Db 227 GlnMetLysLeuPheValAspThrAspAlaAspThrArgLeuSerArgArgGlyIleMet 246
Qy 514 CGGACGCTGCCCGGAGGAGGACCTGAGCAGATT----- 549
Db 247 AsnLeuLeuLeuHisProArgGlyLeuArgAlaIleThrIleAlaValPheGlyLysGln 266
Qy 550 -----CTGACGCAGTACACCACTTCGTGAAGCCGCC----- 582
Db 267 AsnThrTyrIleArgLeuGlnProPheArgIleAsnValProProThrIleThrLysHis 286
Qy 583 TTCGAGAGATTC---TGCCTGCCG----- 603
Db 287 IleGlnGlnLeuGlnCysAlaProValValThrGlnLeuSerArgLysGlnGluThrVal 306
Qy 604 ACAAGAAGTATGCCGATGTGATCATCCACGAGAGAGTG 642
Db 307 ThrValTyrProAlaAspValValLeuPheGlnGlyIle 319
```

Search completed: November 25, 2003, 07:30:45
Job time : 44.5818 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 07:58:14 ; Search time 34 Seconds

(Without alignments)
2102.370 Million cell updates/sec

Title: US-09-896-522-2

Sequence: 1 MASAGEDCSPPADRP.....HPGMLTSGKSHLESSSRH 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932	64.3	261	11	Q8C476
2	873	60.2	201	4	Q96B00
3	733	50.6	260	5	Q8MRJ1
4	570.5	39.3	486	10	Q9FKS0
5	568	39.2	465	10	Q9LTV6
6	563	38.8	483	10	Q9LTK3
7	559	38.6	515	5	Q19583
8	559	38.6	555	5	Q9U317
9	544	37.5	466	10	Q8VYB2
10	542	37.4	614	5	Q9VB10
11	540.5	37.3	626	5	Q8MOK4
12	529	36.5	419	4	Q8N524
13	507.5	35.0	469	10	Q8LD95
14	506	34.9	542	10	Q9L632
15	459	31.7	1060	10	Q9LFF2
16			209	16	Q8E0A8

17	459	31.7	454	3	Q74427	Q74427 schizosacch
18	458	31.6	209	16	Q8E5Y5	Q8E5Y5 streptococc
19	451	31.1	207	16	Q8CSB2	Q8CSB2 streptococc
20	441	30.4	209	16	Q8DTG1	Q8DTG1 streptococc
21	439	30.3	419	5	Q9BMX4	Q9BMX4 cryptospori
22	404	27.9	210	16	Q8EPT5	Q8EPT5 oceanobacti
23	403	27.8	213	16	Q8B8F4	Q8B8F4 vibrio vuln
24	390.5	26.9	111	4	Q9BU42	Q9BU42 homo sapien
25	380.5	26.2	111	4	Q92528	Q92528 homo sapien
26	369	25.4	105	11	Q9QYX8	Q9QYX8 rattus norv
27	349	24.1	212	16	Q8EDX4	Q8EDX4 shewanella
28	233	16.1	674	10	Q9C664	Q9C664 arabidopsis
29	212	14.6	643	10	Q9C9B9	Q9C9B9 arabidopsis
30	210.5	14.5	334	16	Q8EPR9	Q8EPR9 anabaena sp
31	210.5	14.5	403	10	Q8GRU9	Q8GRU9 oryza sativ
32	207	14.3	646	10	Q8S3R9	Q8S3R9 oryza sativ
33	206.5	14.2	396	10	Q95033	Q95033 odontella s
34	200	13.8	334	2	Q8L0S5	Q8L0S5 synchococc
35	200	13.8	334	16	Q8DHN2	Q8DHN2 synchococc
36	199.5	13.8	352	10	P93681	P93681 pieism sativ
37	197	13.6	367	3	Q94642	Q94642 schizosacch
38	196.5	13.6	448	10	Q8GUE1	Q8GUE1 galidieria s
39	195.5	13.5	333	2	Q9L8V7	Q9L8V7 synchococc
40	191	13.2	313	16	Q8YU6	Q8YU6 anabaena sp
41	187.5	12.9	271	10	Q9SPH7	Q9SPH7 beta vulgar
42	178.5	12.3	232	3	Q12084	Q12084 saccharomyc
43	172	11.9	555	16	Q9WZL0	Q9WZL0 thermocoga
44	169.5	11.7	405	10	Q9ATC3	Q9ATC3 vaucheria l
45	156.5	10.8	307	16	Q8DD30	Q8DD30 vibrio vuln

ALIGNMENTS

RESULT 1

Q8C476 PRELIMINARY; PRT; 261 AA.

AC Q8C476; 01-MAR-2003 (T-EMBLrel. 23, Created)

DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)

DE Uridine-cytidine kinase 2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK082837; BAC38646.1; -

SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Query Match 64.3%; Score 932; DB 11; Length 261;

Best Local Similarity 68.8%; Pred. No. 1.4e-75;

Matches 187; Conservative 28; Mismatches 41; Indels 16; Gaps 4;

QY 8 DCSPAPADRP...QRPFLIGVGGTAGSKTVCEKIMELQNEVEQRQRKVVIIQDR 66

DB 4 DSEQTLQNHQPNNGSPFLIGVSGTAGSKSVCAKIYQLQNEVDYHQXVVIIISQDS 63

QY 67 FKVVLTAQKAKALAGGVNFDHDAFPNDIMHRTKXNIVEGKTVVPTVDFTHSRPT 126

DB 64 FKVLTSEQKALAKGVNFDHDAFPNDIMHRTKXNIVEGKTVVPTVDFTHSRPT 123

QY 127 TVVYPADVVLFEGLIVFYSOEIRDMFHLRLFVDTSDVRLSRVRVDV...RRGRDLEQLIT 185

DB 124 VVIYPADVVLFEGLIVFYSOEIRDMFHLRLFVDTSDVRLSRVRVDV...RRGRDLEQLIT 183

QY 186 QYTFVPAPEEFCLPTKKVADVIIIRGVDMVAIINLVHIOIINGDICKTHRGSSNG 245
 DB 184 QYIFVPAPEEFCLPTKKVADVIIIRGVDMVAIINLVHIOIINGDICKTHRGSSNG 232
 QY 246 RSYKRTSEPDHGMILTSKRSLSSESRPH 277
 DB 233 LSKROT--NGYFNGYPRSRGQASSESSSRPH 261

RESULT 2

Q96BJ0 PRELIMINARY; PRT; 201 AA.
 AC Q96BJ0; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Similar to uridine-cytidine kinase 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Straubeberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015547; AAH1547.1; -.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0986; URIDINKINASE.
 KM Kinase.
 SQ SEQUENCE 201 AA; 22790 MW; 0B5F2F00FF7B5363 CRC64;

Query Match 60.2%; Score 873; DB 4; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e-70;
 Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGGEDCSAPADRPORPPLIGVSGTASGKSTVCEKIMELLGQNEVEQORRV 60
 DB 1 MASAGGEDCSAPADRPORPPLIGVSGTASGKSTVCEKIMELLGQNEVEQORRV 60
 QY 61 ILSODRFKVLTAOKAKALGQVNPDPDAPDNDLMHRTLKNVGEKTVVPTVDVTH 120
 DB 61 ILSODRFKVLTAOKAKALGQVNPDPDAPDNDLMHRTLKNVGEKTVVPTVDVTH 120
 QY 121 SRLPETHVVPADVLPFEGILVFYSQEIADMFLRLFVDTSDVRLSRR 169
 DB 121 SRLPETHVVPADVLPFEGILVFYSQEIADMFLRLFVDTSDVRLSRR 169

RESULT 3

Q8MRJ1 PRELIMINARY; PRT; 260 AA.
 AC Q8MRJ1; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE LDI3909P.
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Archirodota; Hexapoda; Insecta; Preygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnak D., Farfan D., Fiske E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.; (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY119583; AAM50237.1; -.
 DR FlyBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0986; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A1E5 CRC64;

Query Match 50.6%; Score 733; DB 5; Length 260;
 Best Local Similarity 67.3%; Pred. No. 1.1e-57;
 Matches 142; Conservative 32; Mismatches 35; Indels 2; Gaps 2;

QY 23 PFLIGVSGTASGKSTVCEKIMELLGQNEVEQORRVYLSODRFKVLTAOKAKALG 82
 DB 27 PFLIGVAGGTASGKSTVCEKIMELLGQNEVEQORRVYLSODRFKVLTAOKAKALG 86
 QY 83 QVNFDPDAPDNDLMHRTLKNVGEKTVVPTVDVTHS-RLEPETHVVPADVLPFEGIL 141
 DB 87 LFNPDHPDAPDNDLMHRTLKNVGEKTVVPTVDVTHS-RLEPETHVVPADVLPFEGIL 146
 QY 142 VFYSQEIADMFLRLFVDTSDVRLSRRVLDVY-RRGRDLQIITQYTFVPAPEEFCL 200
 DB 147 VFYFPRKIRELFHMKLTFVDTSDVRLSRRVLDVY-RRGRDLQIITQYTFVPAPEEFCL 206
 QY 201 PTKKADVIIIRGVDMVAIINLVHIOIINGDICKTHRGSSNG 231
 DB 207 PTKKADVIIIRGVDMVAIINLVHIOIINGDICKTHRGSSNG 237

RESULT 4

Q9FKS0 PRELIMINARY; PRT; 486 AA.
 AC Q9FKS0; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Uridine kinase-like protein.
 GN AT5G40870.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Colombia;
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:131-145(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
 RA Palm C.J., Shin P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB011477; BAB11349.1; -.
 DR HSSP; Q26598; AAN86169.1; -.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000836; PRKtransferase.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00156; Pribosyltran; 1.
 DR Pfam; PF00485; PRK; 1.

Db 229 DFLVPSKKYADVITPRGGDNHVAVDLIVQHTLGGHDLCKY---PNVFIETTTQIR 285
 QY 256 GDH 258
 Db 286 GMH 288

RESULT 7

Q19583 PRELIMINARY; PRT; 515 AA.
 AC Q19583;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE F19B6.1a protein.
 GN F19B6.1 OR F19B6.1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
 OC Rhabdilitidae; Peloderinae; Caenorhabdilitis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93459.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1a; CE05666.
 DR InterPro; IPR006083; PRK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 515 AA; 58452 MW; 9584D947A2D3B5C5 CRC64;

Query Match 38.6%; Score 559; DB 5; Length 515;
 Best Local Similarity 52.9%; Pred. No. 1.1e-41;
 Matches 110; Conservative 44; Mismatches 46; Indels 8; Gaps 2;

QY 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEQRQRKVITLSQDRFYKVLTAEQAKAL 80
 Db 62 KHPFVIGVCGGSASGKTTVAEKIVERLG-----IPWTTILSMDSFRKVLTPPEIKAAH 114
 QY 81 KGOYNFDPAPFNDLMHRTLKNIVEGKTVFPTVDFTVHSRLPETTVVPADVLFEGI 140
 Db 115 ESRYNFDPGNAPFDPDLVEYLKRLREGKSVDPVYDFNTHSRDPSNMKMYGADVLFIEGI 174
 QY 141 LVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVR-RGRDLEQILTQYTTVKAFAEFEC 199
 Db 175 LAFHDERIKNLMKMKVFVDTGDLRLARRIVRVDVTRGRDIDGIMEQYFFVKAFAFDKYI 234
 QY 200 LPTKRYADVITPRGVDNMVAINLIYQHI 227
 Db 235 APCMDSADLIVPRGENDVAIDMIVQNV 262

RESULT 8

Q09317 PRELIMINARY; PRT; 555 AA.
 AC Q09317;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE F19B6.1b protein.

GN F19B6.1 OR F19B6.1B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
 OC Rhabdilitidae; Peloderinae; Caenorhabdilitis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93459.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1b; CE20700.
 DR InterPro; IPR006083; PRK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 555 AA; 62673 MW; D0786FAB98B8CF98 CRC64;

Query Match 38.6%; Score 559; DB 5; Length 555;
 Best Local Similarity 52.9%; Pred. No. 1.3e-41;
 Matches 110; Conservative 44; Mismatches 46; Indels 8; Gaps 2;

QY 21 QRPFLIGSGGTASGKSTVCEKIMELLGQNEVEQRQRKVITLSQDRFYKVLTAEQAKAL 80
 Db 102 KHPFVIGVCGGSASGKTTVAEKIVERLG-----IPWTTILSMDSFRKVLTPPEIKAAH 154
 QY 81 KGOYNFDPAPFNDLMHRTLKNIVEGKTVFPTVDFTVHSRLPETTVVPADVLFEGI 140
 Db 155 ESRYNFDPGNAPFDPDLVEYLKRLREGKSVDPVYDFNTHSRDPSNMKMYGADVLFIEGI 214
 QY 141 LVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVR-RGRDLEQILTQYTTVKAFAEFEC 199
 Db 215 LAFHDERIKNLMKMKVFVDTGDLRLARRIVRVDVTRGRDIDGIMEQYFFVKAFAFDKYI 274
 QY 200 LPTKRYADVITPRGVDNMVAINLIYQHI 227
 Db 275 APCMDSADLIVPRGENDVAIDMIVQNV 302

RESULT 9

Q09YB2 PRELIMINARY; PRT; 466 AA.
 AC Q09YB2;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Putative uracil phosphoribosyltransferase.
 GN AtUGS5810.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Tortumli M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shimizu P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;

"Full Length cDNA of gene At1g55810 (GI:15222778).";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
 RA Lee J.M., Quach H.L., Tang C., Tortumi M., Wu H.C., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis Open Reading Frame (ORF) Clones";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY022218; AAL60039.1; -;
 DR EMBL: AY129346; AAM67479.1; -;
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRFAMs: TIGR00235; udk; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 466 AA; 52443 MW; 3FA6783CB49727F7 CRC64;
 Query Match 37.5%; Score 544; DB 10; Length 466;
 Best Local Similarity 43.6%; Pred. No. 2.2e-40;
 Matches 116; Conservative 52; Mismatches 78; Indels 20; Gaps 8;
 QY 11 SPAP--ADPRQRPPLIVSGTSGKSTVCCKIMELIGQNEVQORQKVVILSDRF 67
 DB 29 SNRPGMAEEREHGGPFLVIGAGAAAGKTTVCMDIMQQL-----HQRRAVVVNDSTR 81
 QY 68 YKVLTAEOKAKKALGQVNFDPDPAFNDMLHRTLKNVEKTVETPPYDVTSH--LPE 125
 DB 82 YHNVNEVELRV--HDVNFHPDPFDTEQLSSMEKRRKQAVDIEYDKSKTKNNVFP 139
 QY 126 TTVVYADVVLFEGILVFSQELRDMFHLRFVDTDSVLSRRVLAD-VRRGRDLEQIL 184
 DB 140 RR-VNSDSVILIEGILIFHDPRVRLDMNKKIFVDADADVLRARIKRDYTEKRDIAVL 198
 QY 185 TQYTTVPKPAFEFECPLTKYADVILPRGVDNVAIYLQIHQDIL-NGDICKMERGG 243
 DB 199 DQSKVEKVPKPEDFIIPTKKYADIIIPRGDNNVAIDLIVQIHITKLGHDLKLY---P 255
 QY 244 NGRSVKRTPESEDPHGMISGKRSH 269
 DB 256 NLVYIOSTFOIRGMHTLIRDSKTKKH 281
 RESULT 10
 Q9V810 PRELIMINARY; PRT; 614 AA.
 AC Q9V810; Q9V810;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE CG4798 protein.
 GN CG4798.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.B., Garg N.S., Galburt W.M., Glasser K.,
 RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kaskel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishiina N.V., Mobarry C., Morris J., Mostrelti A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Snie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye Y., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- ALTERNATIVE PRODUCTS; 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 DR EMBL: AE003803; AAF57873.1; -;
 DR EMBL: AE003803; AAF57873.1; -;
 DR HSSP: Q26998; 1BD3.
 DR Flybase: FBgn0034213; CG4798.
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRFAMs: TIGR00235; udk; 1.
 KM Alternative splicing; Hypothetical protein.
 FT VARSPLIC 1 207 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 614 AA; 68704 MW; C81B6B39DB392A1 CRC64;
 Query Match 37.4%; Score 542; DB 5; Length 614;
 Best Local Similarity 51.5%; Pred. No. 4.8e-40;
 Matches 106; Conservative 42; Mismatches 50; Indels 8; Gaps 3;
 QY 23 PFLIGSGTAGSKSTVCCKIMELIGQNEVQORQKVVILSDRPFKVTATGAKALKG 82
 DB 174 PFLVIGCGSGASAGKTVAAEKIIESL---DVFW---VTLISMDCFKXINEXQHEBALIN 226
 QY 83 QVNFHPDPAFNDMLHRTLKNVEKTVETPPYDVTSHRLDETTVVYADVVLFEGILV 142
 DB 227 ETVFHPDPAFNDMLHRTLKNVEKTVETPPYDVTSHRLDETTVVYADVVLFEGILV 286
 QY 143 FTSQELRDMFHLRFVDTDSVLSRRVLADV-RRGRDLEQILTQYTTVPKPAFEFECPL 201
 DB 287 FTSPEVLKLLDMKIFVDTDPDIRLARLRDLSQGRDLKGVLYKQYLVNWKSYCNYIAP 346
 QY 202 TKKYADVILPRGVDNVAIYLQIHQDIL 227
 DB 347 TWAHADIIVRGDDNNVAIYLQIHQV 372
 RESULT 11
 Q8MOK4 PRELIMINARY; PRT; 626 AA.
 AC Q8MOK4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)


```

Db      99 -HEYNFDHDPADFTHELLSCMEKLRQGAQVADIPKDYFKTY-RSSVFRVNPVTVIIIEGI 156
Qy      141 LVFYSOEIDRMFHLRLFVDTDSVRLSRRLRD-VRRGRLEQILTOYTFVVPAREBFC 139
Db      157 LFFHDRVRLKMMKLFVCTDADVRLARRIKRDTVENGRIQVLOYSKFXVAPDDF1 216
Qy      200 LPTKKYADVLPFGVNDMMVAINLIVOHIDIL-NGDICK 237
Db      217 LPTKKYADIIIPRGDNNHVALIDLIVGHICRKGQHDICK 255

RESULT 14
Q9LGF32 PRELIMINARY; PRT; 542 AA.
ID      09LGF32;
AC      09LGF32;
DT      01-OCT-2000 (TEMBLrel. 15, Created)
DT      01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT      01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE      F4J16.5.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Burkayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      SEQUENCE FROM N.A.
RA      Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA      Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,
RA      Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA      Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
RA      Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA      Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA      Ecker J.R.;
RT      "Genomic sequence for Arabidopsis thaliana BAC F14J16 from chromosome
RT      1."
RN      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      Ecker J.R.;
RN      Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      SEQUENCE FROM N.A.
RA      Chouk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA      Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,
RA      Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA      Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N.,
RA      Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA      Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA      Theologis A., Ecker J.;
RN      Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN      EMBL; ACC002304; AAF79310.1; -.
RN      HSP; Q26996; 1BD3.
RN      InterPro; IPR006083; PRK URK.
RN      InterPro; IPR000764; Uridine_kin.
RN      Pfam; PF00485; PRK.1.
RN      PRINTS; PRO0988; URIDINKINASE.
RN      SEQUENCE 542 AA; 61022 MW; 35427303EB7F3942 CRC64;

```

```

Qy      11 SPAPE---ADRPORPELI-----GVSGCPASGKSTVCEKIMELGQNEVEQRRKVILS 63
Db      29 SNRPBQMAEEBHGQFFVVGESRGVAGNAGKTYVCDNIMQDL-----HDQRLVVA 81
Qy      64 ODREYKVLTAEOKAKALKQVNFDPDADPNDIMHRTLNIVEGKTEVPTVDFVHSR- 122
Db      82 QDSFHNNVNEVELVR--HDVNFDPDADFTQLLSMEKLRGQAVDLPNDFFSKYKN 139
Qy      123 -LPETTVYPADVLPFEGILVYSQETIRDMFHLRFVDT----- 160
Db      140 VFEPFR-VNPSDVIIIEGILIFHDPFRDMMKIFVDAGLSHTKPEVNTYVKSAYAMR 198
Qy      161 -----DSVRLSRVRD-VRRGRLEQILTOYTFVVPAREBFCPTKKYADVLP 211
Db      199 CTCICTHEDADVRLARRIKRDTVENGRIQVLOYSKFXVAPDDF1LPTKKYADIIIP 258
Qy      212 RGVNDMMVAINLIVOHIDIL-NGDICKMHRGSGNGSRKRTSESPDHPGMLTSGKRSH 269
Db      259 RGDNNHVALIDLIVGHHTLGLGQHDICKIY---PNLVIGSTPQIRGMHLLINDSKTTKH 314

RESULT 15
Q9LFF22 PRELIMINARY; PRT; 1060 AA.
ID      09LFF22;
AC      09LFF22;
DT      01-OCT-2000 (TEMBLrel. 15, Created)
DT      01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT      01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE      F20N2.19.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Burkayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      SEQUENCE FROM N.A.
RA      Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,
RA      Shim P., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,
RA      Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA      Lenz C., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
RA      Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA      Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA      Ecker J.R.;
RT      "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome
RT      1."
RN      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      Ecker J.R.;
RN      Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Walker P.P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,
RA      Waliker M.P., Altafi H.H., Araujo R.R., Conn L.L., Conway A.A.B.,
RA      Gonzales A.A., Hansen N.N.F., Huizar L.L., Kremenevskata I.I.,
RA      Lenz C.C., Li J.J., Liu S.S., Lueros S.S., Rowley D.D., Schwartz J.J.,
RA      Toriumi M.M., Vysotskaia V.V., Yu G.G., Davis R.R.W.,
RA      Federspiel N.N.A., Theologis A.A., Ecker J.J.R.;
RN      Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RA      Ecker J.R.;
RN      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      SEQUENCE FROM N.A.
RA      Chouk R., Shim P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA      Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E.,
RA      Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,

```

```

Query Match      35.0%; Score 507.5; DB 10; Length 542;
Best Local Similarity 38.8%; Pred. No. 5, Le-37;
Matches 116; Conservative 52; Mismatches 78; Indels 53; Gaps 10;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 07:59:59 ; Search time 12 Seconds

(without alignments)
1085.533 million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450
Sequence: 1 MASAGEDCSPPAEPADPH.....HPGWLTSGRSHLESSSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	100.0	277	1 UCK1_HUMAN	Q9HA47 homo sapien
2	1325	91.4	277	1 UCK1_MOUSE	P52623 mus musculu
3	934	64.4	261	1 UCK2_HUMAN	Q9BZX2 homo sapien
4	932	64.3	261	1 UCK2_MOUSE	Q9BPM9 mus musculu
5	738	50.9	260	1 UCK_DROME	Q9VCE9 drosophila
6	545.5	37.6	548	1 URL1_HUMAN	Q9HW35 homo sapien
7	539.5	37.2	548	1 URL1_MOUSE	Q9IY13 mus musculu
8	512.5	35.3	229	1 UCK_CABERL	Q17413 caenorhabdi
9	473	32.6	211	1 URK_BACHD	Q9K488 bacillus ha
10	468	32.3	208	1 URK_CLOPE	Q8XJ16 clostridium
11	454	31.3	206	1 URK_LACIA	Q9CF11 lactococcus
12	447	30.8	207	1 URK_STAM	Q92T18 staphylococ
13	439.5	30.3	208	1 URK_STRP8	Q9P018 streptococc
14	439.5	30.3	208	1 URK_STRP8	Q9P018 streptococc
15	438	30.2	212	1 URK_STRP8	Q9P018 streptococc
16	437	30.1	501	1 URK1_YEAST	P27515 saccharomyc
17	430	29.7	211	1 URK_BACSU	Q12033 bacillus su
18	417	28.8	225	1 URK_HALT1	Q9HGC9 halobacteri
19	412	28.4	222	1 URK_CHLNP	Q92710 chlamydia p
20	410.5	28.3	205	1 URK_BORBU	Q9S190 borrelia bu
21	408	28.1	210	1 URK_DEIRA	Q9RX55 diennococcu
22	406	28.0	213	1 URK_VIBCH	Q9K677 vibrio chol
23	398	27.4	209	1 URK_LISIN	Q92B16 listeria in
24	392	27.0	209	1 URK_LISMO	Q9H777 listeria in
25	384	26.5	213	1 URK_HAELN	P44533 haemophilus
26	377.5	26.0	213	1 URK_SALTY	Q8XEY2 salmonella
27	375.5	25.9	213	1 URK_ECOLI	P12128 escherichia
28	371.5	25.6	213	1 URK_MYCEN	P15217 mycoplasma
29	363.5	25.1	213	1 URK_MYCCE	P47622 mycoplasma
30	362.5	25.0	213	1 URK_YERPE	Q8ZT55 yersinia pe
31	362	25.0	216	1 URK_PASMU	Q9CMF5 pasteurella
32	328.5	22.7	207	1 URK_UREPA	Q9P619 ureaplasma
33	217	15.0	404	1 KPPR_WHEAT	P26302 triticum ae

34	205.5	14.2	395	1 KPPR_ARATH	P25697 arabidopsis
35	203	14.0	332	1 KPPR_SYNY3	P37101 synecchocyst
36	200	13.8	397	1 KPPR_MESCR	P27774 mesembryant
37	195.5	13.5	402	1 KPPR_SPIOL	P09559 spiracle ol
38	183	12.6	322	1 COAA_AGRIS	Q8UJ92 agrobacteri
39	183	12.6	375	1 KPPR_CHLRE	P19824 chlamydomon
40	180.5	12.4	319	1 COAA_BACSU	P54536 bacillus su
41	178	12.3	322	1 COAA_BRUME	Q8VE39 brucella me
42	175.5	12.1	306	1 COAA_LISMO	Q8Y810 listeria mo
43	171	11.8	316	1 COAA_PASMU	P57967 pasteurella
44	170.5	11.6	306	1 COAA_LISIN	Q92B16 listeria in
45	166	11.4	331	1 COAA_RHIME	Q92CB5 rhizobium m

ALIGNMENTS

RESULT 1
ID UCK1_HUMAN STANDARD; PRT; 277 AA.
AC Q9HA47;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases.";
RL Mol. Pharmacol. 59:1181-1186(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Ho Y.S., Johnson R.K.;
RT "Human uridine kinase from prostate cancer cell line (LNCap).";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Ilogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Nagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K., Nishimura Y., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y., NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xin Y.R., Yu L., Zhao S.Y.;
RT "Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocyridine, 5-methylcyridine, and N(4)-amsoylcyridine.
CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -|- PATHWAY: Pyrimidine salvage pathway.
CC -|- TISSUE SPECIFICITY: Ubiquitous.

```

CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF237290; AAK28324.1; -
DR EMBL; AF254133; AAK49122.1; -
DR EMBL; AK022317; BAB14010.1; -
DR EMBL; AF125106; AAL75943.1; -
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR000764; uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DCESPAPBAD -> GARAPAGAN (IN REF. 4).
FT CONFLICT 56 57 OR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MW; AFD9ED92780CD502 CRC64;

Query Match 100.0%; Score 1450; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 5,7e-113;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGDECSAPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEOROKRV 60
DB 1 MASAGDECSAPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEOROKRV 60
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVEGKTEVPTDYFVTH 120
DB 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVEGKTEVPTDYFVTH 120
QY 121 SRLPETTVVPADVLFEGILVFYSQEIIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETTVVPADVLFEGILVFYSQEIIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQILTYTTFYKPAFEBCFLPTKKYADYIIRGVNMAINLIYOHIDILNGDICKMHR 240
DB 181 EQILTYTTFYKPAFEBCFLPTKKYADYIIRGVNMAINLIYOHIDILNGDICKMHR 240
QY 241 GGSNGRSYKRTFSEPGDHGMLTSGKSHLSSSRPH 277
DB 241 GGSNGRSYKRTFSEPGDHGMLTSGKSHLSSSRPH 277

RESULT 2
UC1_MOUSE STANDARD; PRT; 277 AA.
ID UC1_MOUSE
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UC1 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UC1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Pange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Halton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97108719; PubMed=8951040;
RA Ropp P.A., Traut T.W.;
RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT brain."
RL Arch. Biochem. Biophys. 336:105-112(1996).
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC025146; AAK25146.1; -
DR EMBL; L31783; AAB50568.1; -
DR MGD; MGI:98904; Umpk.
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
SQ SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Query Match 91.4%; Score 1325; DB 1; Length 277;
Best Local Similarity 91.7%; Pred. No. 1.3e-102;
Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MASAGDECSAPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEOROKRV 60
DB 1 MASAGGGSSAPAPADRPORPFLIGVSGTASGKTVCEKIMELLGONEVEOROKRV 60
QY 61 ILSODRFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVEGKTEVPTDYFVTH 120
DB 61 ILSODCFYKVLTAEOAKALKGOYNFDPADFPNDLMHRTLKNIVEGKTEVPTDYFVTH 120
QY 121 SRLPETTVVPADVLFEGILVFYSQEIIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETTVVPADVLFEGILVFYSQEIIRDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180

```

Oy		181	TQYTFPKPAEEECCLPFXKADVTIPGVNMVAINLIVOHIOIDILNGDICKRR	240
Dd		181	QTAFAPKAFAEEFCPLPXKADVTIPGVNMVAINLIVQHIOIDLNGDLCKRR	240
Oy		241	GGSNGRSYKRTFSEPDHPGMLTSGKRSHLESSSRPH	277
Dd		241	GGPNGRNHKRTFFPEBPGDHGVLATGKRSHLESSSRPH	277
 RESULT 3				
ID	U_CK2_HUMAN	STANDARD:	PRT;	261 AA.
AC	09BXZ2; 096KG5;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (uridine monophosphokinase 2) (Cytidine monophosphokinase 2).			
GN	UCK2 OR UMPK.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=21203613; PubMed=11306702;			
RA	Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;			
RT	"Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases."			
RL	Mol. Pharmacol. 59:1181-1186(2001).			
RN	[2]			
RP	SEQUENCE OF 15-261 FROM N.A.			
RC	TISSUE=Fibrosarcoma;			
RX	MEDLINE=21385121; PubMed=11494055;			
RA	Kojima K., Shimamoto Y., Azuma A., Wataya Y., Matsuda A., Sasaki T.,			
RT	Fukushima M.;			
RT	"Cloning and expression of uridine/cytidine kinase cDNA from human fibrosarcoma cells."			
RL	Int. J. Mol. Med. 8:273-278(2001).			
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-methylcytidine, and N(4)-anisoylcytidine.			
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.			
CC	-1- PATHWAY: Pyrimidine salvage pathway.			
CC	-1- TISSUE SPECIFICITY: Expressed in placenta.			
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-stb.ch/announce or send an email to license@isb-stb.ch).			
CC	-----			
DR	EMBL; AF236637; AAK14053.1; -			
DR	EMBL; AB062451; BAB56162.1; -			
DR	Genew; HGNC:12562; UMPK.			
DR	InterPro; IPR006082; PRK.			
DR	InterPro; IPR006083; PRK_UK.			
DR	InterPro; IPR00764; Uridine_kin.			
DR	Pfam; PF00485; PRK_1.			
DR	PRINTS; PRO0478; PHRIDKLINASE.			
DR	PRINTS; PRO0988; PHRIDINKINASE.			
DR	TIGRFAMs; TIGR00235; uck; 1.			
DR	Transferase; Kinase; ATP-binding.			
NP_BIND	27 34 ATP (POTENTIAL).			

```

SQ      SEQUENCE 261 AA; 29239 MM; 717913146091EBRD CRC64;
Query Match      64.4%; Score 934; DB 1; Length 261;
Best Local Similarity 69.5%; Pred. No. 2.9e-70;
Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;

OY      8 DCESPAEADRPH-QREFLIGVSGCTASGKSTVCEKIMELLGONEVEGRORXVILLSDR 66
DB      4 DSEQTLLNQHPNGEPEPLIGVSGCTASGKSSVCAKIVQLLGONEVDRQRQVILLSDS 63
OY      67 FYKVLTAEOKAKALKGQVNFDPHPAFNDMLAHRTLLKIIVEGTEVPEPLYPDVTHSRRLPET 126
DB      64 FYRLVTEGQRKAKLKGQFNPHPAFNDMLIKLKTEIGTKVQIPIYDVSHSRKEET 123
OY      127 TVVYPADVVLFEGLIVFVSEIIRDMFHRLFVDTSDVRLSRRLVADV-RRGRDLEQILT 185
DB      124 VTVPADVVLFEGLIATFVSGEVRDLFQMKLFVDTDADTRLRSRLVRLDISERGRDLEQLTS 183
OY      186 QYTFYVPAFBEFCLPTKKYADVILPRGVDMVAINLVQHIDLIINGDICKMHRGSGS 245
DB      184 QYTFVPAFBEFCLPTKKYADVILPRGADLVAINLVHIDLIING-----G 232
OY      246 RSYKRTFSEPDHFGMLTSGKSHLESSRPH 277
DB      233 PSKRQT--NCGLNGYTPSRKRGQASSESSRPH 261

RESULT 4
UCK2_MOUSE
ID      UCK2_MOUSE STANDARD; PRT; 261 AA.
AC      099PM9;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN      monophosphokinase 2) (Cytidine monophosphokinase 2).
OS      UCK2 OR UMPK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21203813; PubMed=11306702;
RA      Van Rompay A.R., Nords A., Linden K., Johansson M., Karlsson A.;
RT      "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT      human uridine-cytidine kinases.";
RL      Mol. Pharmacol. 59:1181-1186(2001).
CC      -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC      monophosphate and cytidine monophosphate. Does not phosphorylate
CC      deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC      as a phosphate donor. Can also phosphorylate cytidine and uridine
CC      nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC      thiaouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC      benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC      methylcytidine, and N(4)-antioylcytidine (By similarity)".
CC      -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1- PATHWAY: Pyrimidine salvage pathway.
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF236636; AAKR1052.1; -.
DR      MGD; MGI:1931744; UCK2.
DR      InterPro; IPR006082; PRK.
DR      InterPro; IPR006083; PRK URK.

```


[illegible]

RESULT 5	UCL_DROME	STANDARD;	PRT;	260 AA.
ID	UCL_DROME			
AC	Q9V939;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Probable uridine-cytidine kinase (EC 2.7.1.48) (UCL) (Uridine monophosphokinase) (Cytidine monophosphokinase).			
GN	CG6364.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkelley;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisier S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amannikides P.G., Scherer S.E., Li F.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutcliffe G.G., Wotman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chamme M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abbil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Berkson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,			
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,			
RA	de Palocz B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foeler C., Gebrian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibbewam C.,			
RA	Jatani M., Kalush F., Karpen G.H., Ke Z., Kemsion J.A., Ketchum K.A.,			

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lan Z.,
 RA Laoko P., Leiy, Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., Meled M.P., McPherson D.,
 RA Merkulov G., Milehina N.V., Mobarly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye Y., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcement>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 DR EMBL/ AE003747; AAF56274.1; -
 DR FlyBase; FBgn0039179; CG6364.
 DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00988; URIDINKINASE.
 DR Hypothetical protein; Transferase; Kinase; ATP-binding.
 FT NP BIND 34 41 ATP (POTENTIAL).
 SO SEQUENCE 260 AA; 29213 MW; 2BE57423704925E6 CRC64;

RESULT 6		
URL1_HUMAN		
ID	URL1_HUMAN	STANDARD; prt; 548 AA.
AC	O9HWZ5; O9H3Z2;	
DT	28-FEB-2003 (Rel. 41, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Uridine Kinase-like 1.	
GN	URKL1.	
OS	Homo sapiens (Human) .	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kamagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Iwagaki T., Sugano S.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Cowell G.J., Dedman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.B., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharshah M.H., Leverhulme M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McElay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Pratchalingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showstreen R., Sims S.,
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromane A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitelaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 CC
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: AK000524; BAB91230.1; -;
 DR EMBL: AL118506; CAC13497.1; -;
 DR HSSP: Q26998; 1BD3.
 DR GeneW: HGNC:15938; URRL1.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRIDKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRPFAMs: TIGR00235; udk.1.
 KM Transferrase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD303697E02FE7DB CRC64;
 Query Match 37.64; Score 545.5; DB 1; Length 548;
 Best Local Similarity 45.74; Pred. No. 1.1e-37;
 Matches 111; Conservative 45; Mismatches 60; Indels 27; Gaps 5;
 QY 9 CESAP:-----EADRP:-----HOPFLIGVSGTASGKSYCEKIMELGQ 49
 DB 65 CKSEPLFRTSKRTIYAGRPVWNEHGTSKEAFALGGLGSGASGKTYVARIIEAL-- 122

QY 50 NEVEQRQRKRVILISODRFYKVLTAEOKAKALGVNFHDPDAFPNDLMARLTKNIEGKT 109
 DB 123 -DVPW---VLLSMDSFYKVLTEBOQEOAHNNFEDHPDAFPDPLISTLKAKQSGS 177
 QY 110 VEVPTVDFVTHRLPPTTVVYPADVVLFEGLIVFYSOEIRDFHRLFVDTSDVRLSPR 169
 DB 178 VNVPTDFVTHRLPPTTVVYPADVVLFEGLIVFYSOEIRDFHRLFVDTSDVRLSPR 237
 QY 170 VLRDV-RRGRLDLEQILTYTTTVKPAFEEFCPTKXAVDVIIPRGVDNVAIINLYQH 228
 DB 238 LRRDISERGRLDIEGVLYKYNKFKVSEFDQYIOPVRLADIVVPRGSGTVAIINLYQV 297
 QY 229 DIL 231
 DB 298 SQL 300
 RESULT 7
 URL_MOUSE STANDARD; PRT; 548 AA.
 ID URL_MOUSE
 AC 091Y13;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Uridine kinase-like 1.
 GN URKL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton W., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Pangue C.,
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boscar S., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schermer A., Schein J.E., Jones S.J.M., Maitra W.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: BC016535; AAH16535.1; -;
 DR MGD: MGI:1915806; 1110007H10R1K.
 DR InterPro: IPR006082; PRK.
 DR InterPro: IPR006083; PRK URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK.1.
 DR PRINTS: PR00478; PHRIDKINASE.
 DR PRINTS: PR00988; URIDINKINASE.
 DR TIGRPFAMs: TIGR00235; udk.1.

KW Transferrase; Kinase; ATP-binding.
 FT NP BIND 105 112 ATP (POTENTIAL).
 SQ SEQUENCE 548 AA; 60841 MW; 99A5CA2359BD525 CRC64;

Query Match 37.2%; Score 539.5; DB 1; Length 548;
 Best Local Similarity 45.3%; Pred. No. 3.6e-37;
 Matches 110; Conservative 46; Mismatches 60; Indels 27; Gaps 5;

QY 9 CESFAP-----EADRP-----HQRPFLIGSGSTVCEKXEMELLGQ 49
 DB 65 CKSEPPRLRTSKRTTYTAGRPWPVNEHGTSKEAFALIGGSGASGKTVAEMTEAL-- 122
 QY 50 NEVEQGRGVILSDREFYKVLTAOKAKALGCGYFDPDPAPNDMLHRTLKNIIVEGKT 109
 DB 123 -DVEW----VVLMSDSFYKVLTLQOQOQALACNNFNFPDPAPFDLILSTLKLGKGRS 177
 QY 110 VEPYTFDFTHSRLEPPTTVVPADVLFEGILVFSQEIEMFHLRLFVDTSDVRLSRR 169
 DB 178 VQVPIYDFTTSRRKMDKMTLYGANVILFEGIMAFADKLTLELDMKIFVDTSDIRLVRR 237
 QY 170 VLKRV-RRGRDLEQLQYTTFFVKKPAPEEFCLPTKKYADVILPRGVNMVAINLVQHIO 228
 DB 238 LRRDISERGRDIEGVTKVKKFPAFDQYIQTMRADIVPRSGNTVAIDLIVQHVH 297
 QY 229 DIL 231
 DB 298 SQL 300

RESULT 8
 ID UCK CAEEL STANDARD; PRT; 229 AA.
 UCK CAEEL

AC 017413;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN B0001.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
 OC Rhabdilitidae; Pelodierinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol NZ;
 RA Sims M.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z69634; CAA93453.1; -
 DR PIR; T18629; T18629.
 DR WormPep; B0001.4; CE05142.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK_UK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 KW Hypothetical protein; Transferrase; Kinase; ATP-binding.
 FT NP BIND 15 22 ATP (POTENTIAL).
 SQ SEQUENCE 229 AA; 26195 MW; 89CB21596A820FCC CRC64;

Query Match 35.3%; Score 512.5; DB 1; Length 229;
 Best Local Similarity 50.0%; Pred. No. 2e-35;
 Matches 105; Conservative 40; Mismatches 58; Indels 7; Gaps 4;

QY 23 PFLIGSGTASGKSTVCEKIMELGQNEVEQ-RQRKVILSDREFYKVLTAOKAKALK 81
 DB 8 FLIIGVAGGTSCKSKSTVERITENINAMKQSRQDIIVHLSHSFYRELISAEKILARE 67
 QY 82 GQYFDPHPDNDMLHRTLKNIIVEGTYVEVPYDFVTHSRLEPPTTVVPADVLFEGIL 141
 DB 68 GKNFPHPDQINDPDLAETLQNMIDGKVEIPIKYMIT--SSNMGTVVEPAKVIILEGIL 126
 QY 142 VFSQEIEMFHLRLFVDTSDVRLSRRV---LRDVRGRDLEQLQYTTFFVKKPAPEEF 158
 DB 127 LLYDERVKLLSTKLTFVKKASRLNRLATYIRDRAP--LSITQYTFVKKPAPEEF 184
 QY 199 CLPTKKYADVIIPRGVNMVAINLVQHIO 228
 DB 185 CRPTKKYADVIIPRGADNHVATDLIAKNLQ 214

RESULT 9
 ID URK BACHD STANDARD; PRT; 211 AA.
 URK BACHD

AC 09KDB8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN UPK OR BH1275.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 ON NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AP001511; BAB04994.1; -
 DR PIR; C83809; C83809.
 DR HAMAP; MF_00551; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK_UK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 KW Transferrase; Kinase; ATP-binding; Complete proteome.
 FT NP BIND 12 19 ATP (POTENTIAL).
 SQ SEQUENCE 211 AA; 24387 MW; C2AFA2CB0030520B CRC64;

Query Match 32.6%; Score 473; DB 1; Length 211;
 Best Local Similarity 43.3%; Pred. No. 3.4e-32;
 Matches 93; Conservative 50; Mismatches 56; Indels 16; Gaps 4;

QY 21 QRPPLIGVSGGTASGKSTVCEKINELLGONEVEQORQKVVILSODRFYK--VLTAEOKA 77
 Db 3 KRPIIFIGVAGGTGSGKTVAKEIF-----YQFNKESIVLIEQDAYYKQDSQSLSERL 55
 QY 78 KALKGVNPFHPDAPNDLKHRTLKNIIVEGTVVEPTVPTHSRLPETTVVYPADVLF 137
 Db 56 -----QNYNHPHPLAFDDLLIEHLHSLNGCAIEKPYDYQLHTRSNBVLIVEPKDVIIL 110
 QY 138 EGIIVFYSOEIRDMFHLRLFVDTSDVRLSRVLRADV-RGRDLEQILTOYTFVKKPAFE 196
 Db 110 EGIILLEDPRIRRLDIKIYVDTDADVRIRRWVRDRIINERGRTMESVINGLVVKKPMN 169
 QY 197 EFCLEPTKKYADVILIPRGVNMVAINLIVOHIDIL 231
 Db 171 QFIEPTKRYADVILIPGGQNRVAIDLVMTKIRAIL 205

RESULT 10
 URK_CLOPE STANDARD; PRT; 208 AA.

AC 08XJ16;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN UDK OR CPE1770.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / Type A;
 RX MEDLINE=2166373; PubMed=11792842;
 RA Shima T., Ohtani K., Hirakawa H., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayaishi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic flesh-eater";
 RL Proc Natl Acad Sci U.S.A. 99:996-1001(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic pathway.
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP003191; BAB81476.1; -.
 DR HAMAP; MF 00551; -; 1.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TRIGRAMS; TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 FT NP_BIND 11 18
 SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DDEE CRC64;

Query Match 32.3%; Score 468; DB 1; Length 208;

Best Local Similarity 40.9%; Pred. No. 8.6e-32;
 Matches 88; Conservative 55; Mismatches 56; Indels 16; Gaps 4;

QY 21 QRPPLIGVSGGTASGKSTVCEKINELLGONEVEQORQKVVILSODRFYK--VLTAEOKA 77
 Db 2 KRPIIFIGVAGGTGSGKTVAKEIF-----CIAMIDQDSYYKQSHLSMEDRV 54
 QY 78 KALKGVNPFHPDAPNDLKHRTLKNIIVEGTVVEPTVPTHSRLPETTVVYPADVLF 137
 Db 55 KT-----NIDHPAFAFNNLLVSHLSLNGSHIQKESYDPSIHNRLEDITVKEKEIV 109
 QY 138 EGIIVFYSOEIRDMFHLRLFVDTSDVRLSRVLRADV-RGRDLEQILTOYTFVKKPAFE 196
 Db 110 EGIILLEDPRIRRLDIKIYVDTDADVRIRRWVRDRIINERGRTMESVINGLVVKKPMN 169
 QY 197 EFCLEPTKKYADVILIPRGVNMVAINLIVOHIDIL 231
 Db 170 QFTEPTKRYADVILIPGGQNRVAIDLVMTKIRAIL 204

RESULT 11
 URK_LACLA STANDARD; PRT; 206 AA.

AC Q9CF21;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
 GN UDK OR IL1660.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Holstein A., Wincker P., Manger S., Jallion O., Malarme K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus lactis sep. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -1- PATHWAY: Pyrimidine salvage pathway.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE006396; AAK05758.1; -.
 DR PIR; D86832; D86832.
 DR HAMAP; MF 00551; -; 1.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TRIGRAMS; TIGR00235; udk; 1.
 KW Transferase; Kinase; ATP-binding; Complete proteome.
 FT NP_BIND 11 18
 SQ SEQUENCE 206 AA; 23740 MW; 82E38F052037DBEE CRC64;

Query Match 31.3%; Score 454; DB 1; Length 206;
 Best Local Similarity 42.6%; Pred. No. 1.2e-30;
 Matches 92; Conservative 43; Mismatches 65; Indels 16; Gaps 5;

```

QY 21 ORPFLIGVSGGTASGKSTVCCKIMELLGQNEVEQRKRVILSDRFYK---VLTAROKA 77
D 2 KKTLLIGVTGSGASGKSTVSHALTF-SNE-----RIANIEMHSYKQDSHLEFERT 54
QY 78 KALQGVNFDHPDAFNDLHMRTLKNIVEGKTVFVPTVTHSRLEPPTVVYPADVLF 137
D 55 KT-----NYDHPAFNDLHLENKDLKNGKAVFVPTVYASHRSPTITDFPKDVIIVEGIF 109
QY 138 EGIIVFYSQEIIRDMFHLRFPVDTSDVRLSRVLRDV-RRGDLEQLITQYTTVKAFAE 196
D 110 EGIIVLEBERLRLDMDIKI FVDTDDVRIIRIRREDIERGRSLDVSITQYLDAVKPMYH 169
QY 197 EFCLEPTKCYADVITPRGVDNMVAIINLIYOHIDILN 232
D 170 QFIPTKRYADVIIPBGVSNVGVCDIITTKIASILN 205

RESULT 12
URK_STAMP STANDARD; PRT; 207 AA.
ID URK_STAMP
AC 099TN8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase)
GN UDK OR SA141611 OR SA1439 OR MW1561.
OS Staphylococcus aureus (strain MU50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620,
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyana A.,
RA Mitutani U.I.Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hisekawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kakehi M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet 357:1225-1240 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-acquired MRSA.";
RL Lancet 359:1819-1827 (2002).
RP -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -

```

```

DR EMBL; AP004827; BAB95426.1; -
DR PIR; B89943; B89943.
DR HAMAP; MF_00551; 1.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK. 1.
DR PRINTS; PR00478; PHRIELKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 11
SQ SEQUENCE 207 AA; 23505 MW; FD3A6D4E531DB8D CRC64;

Query Match 30.8%; Score 447; DB 1; Length 207;
Best Local Similarity 41.5%; Pred. No. 4.7e-30;
Matches 88; Conservative 50; Mismatches 58; Indels 16; Gaps 4;

QY 25 LIGVSGGTASGKSTVCCKIMELLGQNEVEQRKRVILSDRFYK---KYLTAOKAKALK 81
D 6 IIGIAGSGSGKSTVINEIKNL-----EGSVALLADQYKQKHLPBERLET-- 56
QY 82 GQYNFDPHPDAFNDLHMRTLKNIVEGKTVFVPTVTHSRLEPPTVVYPADVLFEGIL 141
D 57 ---NYDHPAFNDLHLENKDLKNGKAVFVPTVYASHRSPTITDFPKDVIIVEGIF 113
QY 142 VFYSQEIIRDMFHLRFPVDTSDVRLSRVLRDV-RRGDLEQLITQYTTVKAFAEFCFL 200
D 114 ALBNKVLRLDMDVKKIYVDDADRIIRLRDYPKERGRMSDVINQYLSVREPMHQFTE 173
QY 201 PTKKYADVITPRGVDNMVAIINLIYOHIDILN 232
D 174 PTKKYADVITPRGVDNMVAIINLIYOHIDILN 205

RESULT 13
URK_STAMP STANDARD; PRT; 208 AA.
ID URK_STAMP
AC 080P8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine monophosphokinase).
GN UDK OR SBYM18_1380.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Yeasty L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
RP -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -
CC DR EMBL; AP003362; BAB57773.1; -

```

```

CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AE010058; AAL97975.1; -.
DR HAMAP; ME_00551; -.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TRIGRPMs; TIGR00235; udk; 1.
DR Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23850 MW; BIA76ABF2C5AF511 CRC64;

Query Match 30.3%; Score 439.5; DB 1; Length 208;
Best Local Similarity 39.7%; Pred. No. 2e-29;
Matches 87; Conservative 51; Mismatches 64; Indels 17; Gaps 5;

Qy 21 ORPFLIGVSGGTASGKSTVECKIMELLGQNEVEGRQKVVILSDRFYK---VLTAEQKA 77
Db 3 KKPILLIGVTGSGGGKTSVSRALIDSF-----PNARIMTIOHDSYTKDSHMSFEERV 55

Qy 78 KALKGQYNFDPDPFNDLHRTLNKIVEKTVETPTVDSRLPETTVVPADVLF 137
Db 56 KT-----NYDHPPLAFDTPMIOQLKELLAGRPVDIPIDYKKHTRNTTFRQDPQDVII 110

Qy 138 EGIIVFYSQEI RDMFHLRFVDTSDVRLSRVLRD-VRRGRDLEQILTYTTFFVKPAFE 196
Db 111 EGIIVFEDERLRLDMLDKLFVDTDDDIRIRIRIKRDMERGRSLIESIIDQYTSVVKPMYH 170

Qy 197 EFLCPTKKYADVILIPRGVNMVAIINLVQHIDILNGDI 235
Db 171 QFIKPSKRYADIVIPBGVSNVAIDVINSKIASIL-GEV 208

RESULT 14
URK_STRPY STANDARD; PRT; 208 AA.
AC Q99Z70;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR SP1368 OR SPY3_1042 OR SPS0818.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21132684; PubMed=11296296;
RA Ferricetti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Seate S., Stvorov A.N., Kenon S., Lal H.S., Lin S.F.,
RA Qian Y., Uta H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RA Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=21133808; PubMed=12122206;
RA Bares S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlivert P.M., Mueser J.M.;
RA "Genome sequence of a serotype M3 strain of group A Streptococcus:
RA phage-encoded toxins, the high- virulence phenotype, and clone
RA emergence.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Hayashi H., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayaishi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
RT of S. pyogenes SSI-1, SF370 and MGAS832.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006574; AAK34195.1; -.
DR EMBL: AE014157; AAM79649.1; -.
DR EMBL: AP005143; BAC63913.1; -.
DR HAMAP; ME_00551; -.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TRIGRPMs; TIGR00235; udk; 1.
DR Transferrase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 208 AA; 23816 MW; 510007D58B379DB6 CRC64;

Query Match 30.3%; Score 439.5; DB 1; Length 208;
Best Local Similarity 39.7%; Pred. No. 2e-29;
Matches 87; Conservative 51; Mismatches 64; Indels 17; Gaps 5;

Qy 21 ORPFLIGVSGGTASGKSTVECKIMELLGQNEVEGRQKVVILSDRFYK---VLTAEQKA 77
Db 3 KKPILLIGVTGSGGGKTSVSRALIDSF-----PNARIMTIOHDSYTKDSHMSFEERV 55

Qy 78 KALKGQYNFDPDPFNDLHRTLNKIVEKTVETPTVDSRLPETTVVPADVLF 137
Db 56 KT-----NYDHPPLAFDTPMIOQLKELLAGRPVDIPIDYKKHTRNTTFRQDPQDVII 110

Qy 138 EGIIVFYSQEI RDMFHLRFVDTSDVRLSRVLRD-VRRGRDLEQILTYTTFFVKPAFE 196
Db 111 EGIIVFEDERLRLDMLDKLFVDTDDDIRIRIRIKRDMERGRSLIESIIDQYTSVVKPMYH 170

Qy 197 EFLCPTKKYADVILIPRGVNMVAIINLVQHIDILNGDI 235
Db 171 QFIKPSKRYADIVIPBGVSNVAIDVINSKIASIL-GEV 208

RESULT 15
URK_STRPN STANDARD; PRT; 212 AA.
AC Q97Q77;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR SP1208 OR SPRI090.
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```
OX NCBI_TaxID=1313, 171101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tetteijn H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Kouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Gettinger C.,
RA Gilmour R., Glass J.S., Kojia H., Kraft A.R., Lagace R.E.,
RA Leblanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McChren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Uekunas S.R., Kosteck P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB007421; AAK75315.1; -
DR EMBL; AB008482; AAK9893.1; -
DR PIR; A99008; A99008.
DR PIR; B95140; B95140.
DR TIGR; SP1208; -
DR HAMAP; MF_00551; -; 1.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PRO0478; PHRIBLKINASE.
DR PRINTS; PRO0988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; urk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
KW NP_BIND
FT BIND 12 19 ATP (POTENTIAL).
SQ SEQUENCE 212 AA; 24466 MM; 47B405D377C70A8A CRC64;

Query Match 30.2%; Score 438; DB 1; Length 212;
Best Local Similarity 41.1%; Pred. No. 2.7e-29;
Matches 88; Conservative 42; Mismatches 68; Indels 16; Gaps 4;

QY 22 RPFLIGSGTASGKSTVCEKIMELTGNVEQORQKVVILSDRFYK---VLTAEQKAK 78
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 4 RPIITGVGGSGGKTSVSRALTS-----HPPDEKISMTIEHDSYKXDSHLTFEERVK 56
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 ALKQGVNFDHPDADNDIMHRTLTKNIVEGKTVETVYDFVTHSRLLPETTVVYPADVVLFE 138
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 57 T-----NYDHPFAPDPTDLMIEIQIKELLAGRPVDIPTVDYETHTSRSSKTYRQEPQDVFIWE 111
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 139 GILVFYSQELRDMFHLRLFVDTSDSVRLSRVLADV--RGRDLEQLITQYTTFFVKPAFEE 197
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 112 GILVLEDKRLRLDMDIKIFVDTDDVRLIRIKRDMERKRSLSVYNQTLGVAKPMYHQ 171
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 198 FCPLPTKKYADVIRPGVDNMVAINALIVQHIQDIL 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 172 FIESTKRYADIVPEGSNTVAIDLTLTKIAKIL 205
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: November 25, 2003, 08:06:16
Job time : 14 secs

[illegible]

```

RESULT 3
T18629
hypothetical protein B0001.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T18629
R/Sims, M.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z18999
A/Accession: T18629
A/Status: Preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-229 <W18>
A/Cross-references: EMBL:Z69634; PIDN:CAA93453.1; GSPDB:GN00022; CESP:B0001.4
A/Experimental source: clone B0001
C/Genetics:
A/Gene: CESP:B0001.4
A/Map position: 4
A/Introns: 21/3; 41/3; 108/2
C/Superfamily: uridine kinase

Query Match      35.3%; Score 512.5; DB 2; Length 229;
Best Local Similarity 50.0%; Fred. No. 2.1e-35;
Matches 105; Conservative 40; Mismatches 56; Indels 7; Gaps 4;

QY      23 PFLIGSGTASGASTVCEKIMELGQNEVQC-RQKRVILSODRFYVLTAEOGAKALK 81
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      8 PFLIGVAGCTGCGKSTIVERIIEINLNNAKSGRIDIVHLSTLSFRELSAEKKILARE 67
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY      82 GQYNFDPDADFNDIMHRTLKNIVEGKTEVEVPTDFVTHSFLPETTVVYPADVVLFGIL 141
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      68 GKPNFDHDDQINFDLAETLQNMIDGKTVEI:PKYMIT-SSMNGIVTVEPAKVIILIGIL 126
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY      142 VFSGEIRIDMFHLRLFVNTDSDVRLSRV--LRDVRGRRLDQILQYTFVFKPAPPEEF 198
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      127 LLYDNRVAKLLSTKYLEKEKNASRLRNRLATYIRYHRAP--LSIRYTFVFKPAPPEEF 184
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY      199 CLPTKRYADVIIIRGVDDNMVAIINLIVCHIQ 228
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      185 CRPTKRYADVIIIRGADNMVATDLIAKRLQ 214
      |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 4

```

C83809
uridine kinase udk [imported] - *Bacillus halodurans* (strain C-125)
C.Species: *Bacillus halodurans*
C.Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C.Accession: C83809
R.Takami, H.; Nakaone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A.Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A.Reference number: A63650; MUID:20512582; PMID:11058132
A.Accession: C83809
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-211 <STO>
A.Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BAB04994.1; GSPDB:GN0001
A.Experimental source: strain C-125
C.Genetics:
A.Gene: udk
C.Superfamily: uridine kinase

```

Query Match Similarity      32.6%; Score 473; DB 2; Length 211;
Best Local Similarity      43.3%; Pred. No. 3, 7e-32;
Matches      93; Conservative      50; Mismatches      56; Indels      16; Gaps      4;

QY      21 ORPFLIGSGTAGSKSTVCEKIMELLGONEVEQRORKVILISQDRPYK--VLTAEOKA 77
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      3 KRPIIIVAGGTGSCKTVAKEIF-----YQFNKSITVLREQDAYXKQSOLSEERL 55
QY      78 KALKQYNFDHDAFNDNLMHRTLNKIIVEGKTVEVPTYPFTVHSRLPETTVVPADVVLF 137
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      56 ----QTNYVDHLAPDNDLLIELHSLNGOALIEKVYDYKLTRENEVILVEPKVIL 110
QY      138 EGIIVYSOEIRDMFHRLFEVDTSDVRLSRRVLRDPVR-RGRDLBOILLQYTTFVKPAFE 196
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      111 EGILLEBRLEMLDKLFDVIDADIRIRIKRVARIRERGRILSVIEGYTKVVRPMHM 170
QY      197 EFCLPTKKYADVIIIPRGVDMVAINLVIOHIODIL 231
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      171 QPIETKRIVADVIIPEGGONRVAILDMVTIKIRALI 205

RESULT 5
T41020
probable uridine kinase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 03-Dec-1999
C:Accession: T41020
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z21965
A:Accession: T41020
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-454 <MUR>
A:Cross-references: EMBL:AL023860; PIDs:CAA19591.1; GSPDB:GN00068; SPDB:SPCC162.11c
A:Experimental source: strain 97Zh-, cosmid c162
C:Genetics:
A:Gene: SPDB:SPCC162.11c
A:Map position: 3

Query Match      31.7%; Score 459; DB 2; Length 454;
Best Local Similarity      41.3%; Pred. No. 1, 5e-30;
Matches      92; Conservative      47; Mismatches      76; Indels      8; Gaps      2;

QY      11 SPAPEADPHQRPFLIGSGTAGSKSTVCEKIMELLGONEVEQRORKVILISQDRPYKV 70
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      9 SSNPFYEPPMKVRIIGIAGPSGSGKTSVAQLIVKALNPHE-----VILSIDSYKS 61
QY      71 LTAEQAKAKLKQYNFDHDAFNDNLMHRTLNKIIVEGKTVEVPTYPFTVHSRLPETTVV 130
       |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      62 LNAEQKKAFAFNNDYDFDSDEALIDWDLFPVKLETIKQGRKVDIPITYSFNEHRNLPETVTLF 121
QY      131 PADVVLEFGILVFYSOEIRDMFHRLFEVDTSDVRLSRRVLRDPVR-RGRDLBOILLQYTTF 189
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


D_b 122 GASIIILEGI PALYDEKIRSLDVSFLDTSDVCLSRRLNRDINRGRIWGVLEQYNK 181
Q_y 190 FVKPAAEEBCLPYKKADVIIPRGVDNMALIVQHIDILN 232
182 FVKPSTENFVRQLSTDDIIVPRGRDNKAIDMVFNFRRTLIS 224

RESULT 6

uridine kinase (EC 2.7.1.48) [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL14)
 C/Species: *Lactococcus lactis* subsp. *lactis*
 C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C/Accession: D86832
 R/Bolotin, A.; Winkler, P.; Manger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich
 Genome Res. 11, 731-753, 2001
 A/Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
 A/Reference number: A86625; MUID:21235186; PMID:11337471
 A/Accession: D86832
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-206 <SPO>
 A/Cross-references: GB:AA005176; PID:912724672; PIDN:AAK05758.1; GSPDB:GN00146
 A/Experimental source: strain IL1403
 C/Genetics:
 A/Gene: udk
 C/Superfamily: uridine kinase
 C/Keywords: phosphotransferase

Query Match	31.3%	Score 454;	DB 2;	length 206;
Best Local Similarity	42.6%	Pred. No. 1.4e-30;		
Matches 92; Conservative	43;	Mismatches 65;	Indels 16;	Gaps 5

Q3

Db 2 KKLLILGVGGSGAGSTVS^HAILETF-SNE-----RIAMIEHDSYKQOSHLEETBERT 54
Qy 78 KALKGQINFEHPAPFNDLMHRTLNKIYEKGTVKEVPTDFTLHSRLPETTWVPADVLVF 137
Db 55 KT-----NDDPLAFPTDYLIQAOLKELOYGRAVDIPIDYAKHRSQETTRQEVEVDVLIIV 109
Qy 138 EGIIVFYSOEIRDMFHLRLPVDTSBVRLSRRVLADV-RGRSDI^EQLIYYTTFVKFAFE 166
Db 110 EGILVIEDERLRDLMDIKIFVDTDDDVRIIRIRIRDI^EBKRGLDSVTIYLDAVKRMVH 169
Qy 197 EFCLPTKKYADVIIPRGVDNMVAINDLYOHODIIN 232
Db 170 QFIETPKRYADVIIPEGVNTVGAVIIITTKIASIIN 205

RESULT 7

uridine kinase [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89943
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, T.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference numbers: A89758; MUID:21311952; PMID:11418146
A:Accession: B89943
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-207 <KUR>
A:Cross-references: GB:BA000016; PTD:G13701409; PIDN:BAB42703.1; GSEDB:GN00149
A:Experimental source: strain N315

Query Match	30.8%	Score 447	DB 2	Length 2077
Best Local Similarity	41.5%	Pred. No. 5,4e-30		
C:Superfamily: uridine kinase				

Matches	88;	Conservative	50;	Mismatches	58;	Indels	16;	Gaps	4;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

```

25 LIGVSGTASGKSTVCERIMELLGQNEVEQRQRKVIIISODRF--KLTAEQAKALK    81
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
6 IGIAGSGSGKTTVTNEMIKUL-----EGHSVALLAQDYVKDQKHLPDELET--    56

```

```

82  GQNEFHHPDAFENDLMMRTKLLVEGKEVPEVPYDPATSRLEBETVYVPADVLPEGL 147
57  ---NYHPEFPEFDDLLLENKADLKNQKANEVPTTYSKHSRSDITTDKFKKDYIVEGIF 113
142  VPEYSQIEIRDMFHLRFVVDTSDVRLSRRLVADVR-RCGRDEQLLLTOYTFVKAPEEFC 200
114  ALBNKVLRLDMWADYIVYDTPDADRILRLRLTRQKERSMSDVSINOLSVVRPMHDFIE 174

```

201 PTKKYADVIRPGVDNMVAINLIVQHIQDILN 232

174 PTKKYADII IPEGGSNKVAIDIMTTKIQLVS 205

RESULT &

uridine kinase [imported] - Streptococcus pneumoniae (strain TIGR4)
 Species: Streptococcus pneumoniae
 Update: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 Accession: B95140
 RetireIn: H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap
 son, T.; Hickey, E.R.; Holt, I.E.
 Science 293, 498-506, 2001
 Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 Reference number: A95000; MUID:21357209; PMID:11463916
 Accession: B95140
 Status: preliminary
 Molecule type: DNA
 Residues: 1*212 <KUR>
 Cross-references: GB:AE005672; PIDN:AAK75315.1; PID:g149472689; GSFDB:GN00164; TIGR:SP44
 Experimental source: strain TIGR4
 Genetics:
 Gene: SPI208
 Superfamily: uridine kinase

Query match	30.28; Score 438; DB 2; Length 212;
-------------	-------------------------------------

Matches	88;	Conservative	42;	Mismatches	68;	Indels	16;	Gaps	4;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

24

```

4 RPIIIIGTGGGGGKTKTSVRALIS-----HPEDEKISMIIEHDSYKQSHLTFEBERVK 56
79 ALKGQVNFHGHDPADFNDDLMLRHLTKNIVEGKTVEVPYTPDFTHSRLDEPTTVVPADVILFE 138
57 T-----NYDHPFADDDTLMEQIKELAGLPVDIPTYDYEHRRSSKTRQEPQDFVIE 111
139 GILVFYSQEIHRDFHRLFLPVDITSDVRLSRVRADV-RRGRDLQQLITQYTFEVRKAFEE 197
112 GILVLEBKRRRLDMLDKIFVDTDDVRIIRIRKDMEBEERSLDSVINQYLSGVKKPMYHQ 171
198 FCLPTEKKIADVIRPGVDNNVAINLIVQHIQDIL 231
172 FIESTKRVADIVPEGVSNVAIDLTTKAKLL 205

```

RESULT 9

penicillinase (EC2.7.1.48) [imported] - Streptococcus pneumoniae (strain R6)
 Species: Streptococcus pneumoniae
 Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 Accession: A99008
 Authors: Hoskins, J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Eickbush, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001

```

A:Authors: Yang Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Vaskunas, S.R.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: AB7872; MUID:21429245; PMID:11544234
A:Accession: A89008
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <KUN>
A:Cross-references: GB:AE007317; PIDN:AAK9893.1; PID:G15458714; GSPDB:GN00174
C:Genetics:
A:Gene: udk
C:Superfamily: uridine kinase
C:Keywords: phosphotransferase

Query Match      30.2%; Score 438; DB 2; Length 212;
Best Local Similarity 41.1%; Pred. No.3.le-29;
Matches 88; Conservative 42; Mismatches 68; Indels 16; Gaps 4;

QY 22 RPFLIGVSGTASGKSTVCSEKIMELDGGNEVEORQRKVILSDREFYK---VLTAEGQAK 78
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 4 RPIIIGTVGGSGCGKTSVSRAILS-----HFPEKISMTLHDSYYDDQHLLFEEERVK 56
   |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

QY 79 ALKGQYNLDHPDAFPNDLMHRTLKNIIVEGKYVEVTPTFTVSHSLRPETVVYYPADVLIFE 138
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||
Db 57 T-----NNDHPAPFDPTDMIEQIKELLAGRVPDIPTYTYHTTHSSKKYRGEPDVFIIVE 111
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||

QY 139 GILVFYSOEIRDMFEPLRLRVFVDTSDFVRLSRREVLRDV--RRGDLEQLITQYTTFYKPAFEE 197
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||
Db 112 GILVLEDKRLRDLMDIKIFVPTDDPVRITRIKRDMEEGRSLDSVINQYLGVGVPMWYHQ 171
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||

QY 198 FCLPTKKYADVIRPGVNMYALINLYQHIDIL 231
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||
Db 172 FIESTKRYADIVIRGEVSNVAIDLTLTKIAKIL 205
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||

```

RESULT 10
S29374
uridine kinase (EC 2.7.1.48) - Yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein N2050; protein YNR012w
C/Species: *Saccharomyces cerevisiae*
C/Date: 28-May-1993 #sequence revision 28-May-1993 #text_change 23-Mar-2001
C/Accession: S29374; S45134; S48349; S63338; S21361
R/Kern, L.
Nucleic Acids Res. 18, 5279, 1990
A/Title: The UKR1 gene of *Saccharomyces cerevisiae* encoding uridine kinase.
A/Reference number: S29373; MUID:90384830; PMID:2169608
A/Accession: S29374
A/Molecule type: DNA
A/Residues: 1-501 <KER>
A/Cross-references: EMBL:X5398; NID:g4771; PIDN:CAA37946.1; PID:g4773
R/Vershaell, P.; Aert, R.; Voet, M.; Volckaert, G.
Submitted to the EMBL Data Library, January 1994
A/Description: Twelve open reading frames revealed on the 23.6 kbp segment flanking the
A/Reference number: S45119
A/Accession: S45134
A/Molecule type: DNA
A/Residues: 1-501 <VER>
A/Cross-references: EMBL:X77395; NID:g496717; PID:g496728
R/Vershaell, P.; Aert, R.; Voet, M.; Volckaert, G.
Yeast 10, 1355-1361, 1994
A/Title: Twelve open reading frames revealed in the 23.6 kb segment flanking the centromere
A/Reference number: S48338; MUID:95208356; PMID:7900425
A/Accession: S48349
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-501 <VE2>
A/Cross-references: EMBL:X77395; NID:g496717; PIDN:CAA54580.1; PID:g496728
A/Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1994
submitted to the Protein Sequence Database, April 1996
A/Reference number: S62910
A/Accession: S63338
A/Molecule type: DNA
A/Residues: 1-501 <AER>

A:Cross-references: EMBL:771827; NID:g1302489; PID:e239785; PID:g1302490; MIPS:YNR012w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:URK1
A:Cross-references: SGD:S0005295; MIPS:YNR012w
A:Map position: 14R
C:Keywords: phosphotransferase; transmembrane protein
F:154-170/Domain: transmembrane #status predicted <TM1>
F:419-433/Domain: transmembrane #status predicted <TM2>

Query Match	30.1%	Score 437	DB 2	Length 501
Best Local Similarity	38.4%	Pred. No. 1.2e+8		
Matches	93	Conservative	42	Mismatches 75; Indels 32; Gaps 4
QY	19	PHQRPFPLVSGSGTASGKTSVCEKIMELLGQNEVEGRQRKRVVLLSQDRFVKYUQLTAEQAK	78	
DB	52	PPTTYYIIIGIGASGSGKTSVAAKIVSSI-----NVPTVTLISIDNFPNPGDEBRAR	104	
QY	79	ALKGQYNPDPAFDNDLMDHRTLKNIVEKTSVEPTPYDFVTHSRLEP-ITVVVPADVLP	137	
DB	105	AFKNIEYDFDEPANAINDLAVKCIILNKEGKRTNI PYVSFPHNRVRPDKNIVIVGASVYVI	164	
QY	138	ESGIIVFSGEIRDMFHLRLFPVDPDSDVRLSRVLRP-VRGRPLDQILNQYTFPVKPAFE	196	
DB	165	ESIVALYDRRLDLDMDLKIYDADDDVCLARRSRDIVSRGRDLDGCIQOMEKFAKENAV	224	
QY	197	EFCELPFKKXADVIRPGVDNMVAINLIVOH-----QDILNG	233	
DB	225	KFVKRPTMKAAADAIIPMSMDNATAVNLINHIKSKTELKSNHRLRELITKLGSSEQDVILNR	284	
QY	234	DI	235	
DB	285	NI	286	

RESULT 11
 G69728
 uridine kinase udk - Bacillus subtilis
 C|Species: Bacillus subtilis
 C|Date: 05-Dec-1997 #sequence_reviston 05-Dec-1997 #text_change 21-Jul-2000
 C|Accession: G69728
 R|Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.-M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Ch'ang, S.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 380, 249-256, 1997
 A|Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallardo, A.; Harwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koester, P.; Koningsstein, G.; Krogsh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.; Laubert, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee, Y.; Ogawa, K.; Ogikawa, A.; Oudgaa, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Sero, A.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A|Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A|Reference number: A69580; MUID:98044033; PMID:9384377
 A|Accession: G69728
 A|Status: preliminary; nucleic acid sequence not shown; translation not shown
 A|Molecule type: DNA
 A|Residues: 1-211 <KON>
 A|Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14675.1; PID:g2635179
 A|Experimental source: strain 168
 C|Genetics:
 A|Gene: udk
 C|Superfamily: uridine kinase
 Query Match 29.7%; Score 430; DB 2; Length 211;
 Best Local Similarity 39.1%; Pred. NO.1.4e-28;
 Matches 84; Conservative 52; Mismatches 63; Indels 16; Gaps 4;

```
Db 3 KNPVVIAGSGSGKTSVTRSIYE-----QFGKSHILMIQQDLVYKQDLSHPLEBRL 55
Qy 78 KALGQVNFHPDAFNDMLHRTLNKIVEGTVVPTDYPTSHSRLEPTTVVYADVLF 137
Db 56 NT-----NYDHPDAFNDYLLIEHIQDLNYPRIEKPIYDKLHTRSEETVYVPEKDYIL 110
Qy 138 EGIIVFYSQIEIRDMFHRLFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTFFVKPAFE 196
Db 111 EGIIVLEDRKRLMDIRLFDVTDADRLIRIRKIRMDIERKRSIDSVTEQVSVVRPMHN 170
Qy 197 EFCLPTRKYADVILIPRGVDNMVAIINLIVQ--HIQ 231
Db 171 QFVPTPKRYADILIPREGQNHVAIDLMTVKIQITIL 205

RESULT 12
F84277
uridine kinase [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C/Accession: F84277
R/Ref: W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: F84277
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-248 <STO>
A/Cross-references: GB:AE004437; NID:g10580750; PIDN:AA619586.1; GSPDB:GN00138
C/Genetics:
A/Gene: urk
C/Superfamily: uridine kinase

Query Match 28.4%; Score 417; DB 2; Length 248;
Best Local Similarity 39.7%; Pred. No. 2.2e-27;
Matches 83; Conservative 50; Mismatches 58; Indels 18; Gaps 5;

Qy 24 FLIGVSGTASGKSTVCEKIMELLGQNEVEQRQRKRVILSQDRFYKLT--AEQAKAL 80
Db 29 FALIGAGTAGKTTVAKEITDNVGS-----ATLIPLDNYEDLSRPFERRANA- 79
Qy 81 KQGVNFHPDAFNDMLHRTLNKIVEGTVVPTDYPTSHSRLEPTTVVYADVLFEG 140
Db 80 ----NYDHPDAFNDYLLIEHIQDLNYPRIEKPIYDKLHTRSEETVYVPEKDYIL 135
Qy 141 LVFYSQIEIRDMFHRLFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTFFVKPAFE 199
Db 136 LAISDETVNDMLDHIYETDADRLIRIRKIRMDIERKRSIDSVTEQVSVVRPMHN 195
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQ--HIQ 228
Db 196 EPTKRYADILIPREGQNHVAIDLMTVKIQITIL 223

RESULT 13
D86582
uridine kinase [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C/Accession: D86582
R/Ref: R.; Shirai, M.; Hiraoka, H.; Kinoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: D86582
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-222 <STO>
A/Cross-references: GB:BA000008; NID:g8979107; PIDN:BA98942.1; GSPDB:GN00142
```

```
A/Experimental source: strain J138
C/Genetics:
A/Gene: CPJ0735
C/Superfamily: uridine kinase

Query Match 28.4%; Score 412; DB 2; Length 222;
Best Local Similarity 38.4%; Pred. No. 4.9e-27;
Matches 84; Conservative 54; Mismatches 59; Indels 22; Gaps 6;

Qy 25 FLIGVSGTASGKSTVCEKIMELLGQNEVEQRQRKRVILSQDRFYK--VLTAEQAKAL 81
Db 8 FLITGSGAGKTTLNQNIKEIFED-----VAVICQDNYKXRSHTPEERANLI- 58
Qy 82 GQVNFHPDAFNDMLHRTLNKIVEGTVVPTDYPTSHSR-LEPTTVVYADVLFEG 140
Db 59 ----WDHPDAFNDYLLIEHIQDLNYPRIEKPIYDKLHTRSEETVYVPEKDYIL 114
Qy 141 LVFYSQIEIRDMFHRLFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTFFVKPAFE 199
Db 115 LVFENQELRLMDIRLFDVTDADRLIRIRKIRMDIERKRSIDSVTEQVSVVRPMHN 174
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQ--HIQ 234
Db 175 EPTKRYADILIPREGQNHVAIDLMTVKIQITIL 213

RESULT 14
E72041
uridine kinase CP0011 [imported] - Chlamydia pneumoniae (strains CWL029 and AR39)
C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: E72041; E81622
R/Ref: R.; Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: E72041
A/Molecule type: DNA
A/Residues: 1-222 <AN>
A/Cross-references: GB:AE001655; GB:AE001363; NID:g4377039; PIDN:AA018874.1; PID:g4377043
A/Experimental source: strain CWL029
R/Ref: R.; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, J.;
C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: E81622
A/Molecule type: DNA
A/Residues: 1-222 <REA>
A/Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AA637907.1; PID:g7188950
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CP00735; CP0011
C/Superfamily: uridine kinase

Query Match 28.4%; Score 412; DB 2; Length 222;
Best Local Similarity 38.4%; Pred. No. 4.9e-27;
Matches 84; Conservative 54; Mismatches 59; Indels 22; Gaps 6;

Qy 25 FLIGVSGTASGKSTVCEKIMELLGQNEVEQRQRKRVILSQDRFYK--VLTAEQAKAL 81
Db 8 FLITGSGAGKTTLNQNIKEIFED-----VAVICQDNYKXRSHTPEERANLI- 58
Qy 82 GQVNFHPDAFNDMLHRTLNKIVEGTVVPTDYPTSHSR-LEPTTVVYADVLFEG 140
Db 59 ----WDHPDAFNDYLLIEHIQDLNYPRIEKPIYDKLHTRSEETVYVPEKDYIL 114
Qy 141 LVFYSQIEIRDMFHRLFVDTSDVRLSRVLRDV-RRGRDLEQILTYTTFFVKPAFE 199
Db 115 LVFENQELRLMDIRLFDVTDADRLIRIRKIRMDIERKRSIDSVTEQVSVVRPMHN 174
Qy 200 LPTKRYADVILIPRGVDNMVAIINLIVQ--HIQ 234
Db 175 EPTKRYADILIPREGQNHVAIDLMTVKIQITIL 213
```

Db 175 EPTKRYADIIVHGNRYQNVNTIISQIKHLENALESD 213

RESULT 15

G70101

uridine kinase (udk) homolog - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Nov-2000

C:Accession: G70101; T46979

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kervilave, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, W.; Vugt,

Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A:Reference number: A70100; PMID:98065943; PMID:9403685

A:Accession: G70101

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-207 <KLE>

A:Cross-references: GB:AE001115; GB:AE000783; NID:g2687879; PIDN:AA066392.1; PID:g268788

A:Experimental source: strain B31

R:Boursaux-Bude, C.; Margatita, D.; Gilles, A.M.; Barzu, O.; Girons, I.S.

FEMS Microbiol. Lett. 151, 257-261, 1997

A:Title: Borrelia burgdorferi uridine kinase: an enzyme of the pyrimidine salvage pathwa

A:Reference number: 224339; PMID:97372541; PMID:9228761

A:Accession: T46979

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 2-207 <BOU>

A:Cross-references: EMBL:X97449; NID:g1405440; PIDN:CAA66081.1; PID:g1405442

A:Experimental source: strain HB19, sep. senso stricto

C:Genetics:

A:Gene: udk

C:Superfamily: uridine kinase

C:Superfamily: uridine kinase

Query Match 28.3%; Score 410.5; DB 2; Length 207;

Best Local Similarity 36.9%; Pred. No. 6e-27;

Matches 79; Conservative 49; Mismatches 59; Indels 27; Gaps 4;

Db 5 LIGVSGTASGKSTYCEKIMELGONEVEQORQKVVILSDRFYKVLTAEQKAKALGQY 84

5 IIGISGSGSGKSTVWSKISEFIEP-----FVLISQDNYYKSV-----GDY 45

QY 85 -----NFHHPDAFNDLDMRTLKNIVEGKTVSEVPTVDFTHSRLPETTVVVPADVLF 137

Db 46 EHERSKVNFDPDAFDNNLFTEHLKNLKNKSPIDMPLDYDFINHKQQLKTVLVVPTPVVIV 105

QY 138 EGIIVFYSQEIIRDMFHLRLFVDTSDVRLSRVLADV-RRGRDLQILVQYTFEVPKAFE 196

Db 106 EGINIFVEERVRNLDLKIYIDTPNDIFIRRLRDISKRGRTVESVIDQYLNTRMGY 165

QY 197 EFCUPTKRYADVIIPRGVDNVAIIVQHIQDI 230

Db 166 RFIEPTKEYADLIIPGGHNDKALYVLSTFLKSL 199

Search completed: November 25, 2003, 08:07:26

Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:05:10 : Search time 31 Seconds
(without alignments)
1648.090 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450

Sequence: 1 MASAGEDESPAPADRP...HPGMLTSGKSHLESSRPH 277

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 18443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	100.0	277	US-09-896-522-2	Sequence 2, Appl1
2	896	61.8	337	US-09-925-300-1160	Sequence 2, Appl1
3	428.5	29.6	125	US-09-896-522-5	Sequence 5, Appl1
4	327	22.6	60	US-10-029-386-28765	Sequence 28765, A
5	320.5	22.1	124	US-09-896-522-6	Sequence 6, Appl1
6	276	19.0	231	US-09-896-522-4	Sequence 4, Appl1
7	154.5	10.7	312	US-09-738-626-4609	Sequence 4609, A
8	149	10.3	339	US-10-156-761-12495	Sequence 12495, A
9	119.5	8.2	143	US-09-734-569-88	Sequence 88, Appl1
10	97	6.7	216	US-10-301-822-217	Sequence 217, Appl1
11	95	6.6	251	US-10-301-822-220	Sequence 220, Appl1
12	95	6.6	257	US-10-106-698-6128	Sequence 6128, Appl1
13	95	6.6	267	US-10-301-822-215	Sequence 215, Appl1
14	93	6.4	210	US-09-769-744A-64	Sequence 64, Appl1
15	93	6.4	433	US-09-815-242-5691	Sequence 5691, Appl1

16	93	6.4	433	9	US-09-815-242-12690	Sequence 12690, A
17	91	6.3	234	15	US-10-156-761-10887	Sequence 10887, A
18	88	6.1	674	9	US-09-765-272-200	Sequence 200, Appl1
19	87	6.0	457	12	US-09-882-227-394	Sequence 394, Appl1
20	87	6.0	865	9	US-09-815-242-11364	Sequence 11364, A
21	87	6.0	910	12	US-09-896-166B-16	Sequence 16, Appl1
22	86	5.9	457	9	US-09-815-242-11616	Sequence 11616, A
23	84.5	5.8	252	9	US-09-925-301-1273	Sequence 1273, Appl1
24	84.5	5.8	508	9	US-09-815-242-11555	Sequence 11555, A
25	84	5.8	600	11	US-09-893-519A-71	Sequence 71, Appl1
26	82.5	5.7	214	12	US-10-032-585-7414	Sequence 7414, Appl1
27	82.5	5.7	231	12	US-10-220-381-17	Sequence 17, Appl1
28	82	5.7	665	15	US-10-216-556-18	Sequence 18, Appl1
29	81.5	5.6	508	12	US-09-882-227-570	Sequence 570, Appl1
30	81	5.6	213	15	US-10-156-761-13762	Sequence 13762, A
31	81	5.6	245	15	US-10-156-761-11353	Sequence 11353, A
32	80.5	5.6	722	8	US-08-817-832B-32	Sequence 32, Appl1
33	80	5.5	645	10	US-09-919-172-41	Sequence 41, Appl1
34	80	5.5	744	10	US-09-835-081-4	Sequence 4, Appl1
35	80	5.5	877	9	US-09-815-242-4876	Sequence 4876, Appl1
36	80	5.5	880	9	US-09-815-242-10675	Sequence 10675, A
37	79.5	5.5	231	9	US-09-815-242-5760	Sequence 5760, Appl1
38	79.5	5.5	744	9	US-09-815-242-5473	Sequence 5473, Appl1
39	79.5	5.5	744	9	US-09-815-242-12346	Sequence 12346, A
40	79	5.4	1070	12	US-10-032-585-7389	Sequence 7389, Appl1
41	78.5	5.4	414	15	US-10-119-651-6	Sequence 6, Appl1
42	78.5	5.4	610	9	US-09-815-242-10414	Sequence 10414, A
43	78.5	5.4	691	10	US-09-919-585-6	Sequence 6, Appl1
44	78.5	5.4	724	10	US-09-919-585-9	Sequence 9, Appl1
45	78.5	5.4	730	15	US-10-217-371-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-896-522-2
Sequence 2, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 277
TYPE: PRT
ORGANISM: Homo sapiens
US-09-896-522-2

Query Match 100.0%; Score 1450; DB 9; Length 277;
Best Local Similarity 100.0%; Pred. No. 5.7e-151;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 60
DB 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 60
QY 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 120
DB 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 120
QY 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 180
DB 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 180
QY 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 240
DB 1 MASAGEDESPAPADRP...PPLIGVSGTASGKTCCKIMELGONEVEQRKRV 240

Db 181 EQLITQTTTTPKPAPEEFCLPTKXADYIIPRGDNNVALNIQIIDLINGDICKMR 240
QY 241 GGSNGRSYKRTFSPGDHPGMLTSGKSHLESSSRPH 277
Db 241 GGSNGRSYKRTFSPGDHPGMLTSGKSHLESSSRPH 277

RESULT 2
US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Query Match 61.8%; Score 896; DB 10; Length 337;
Best Local Similarity 62.3%; Pred. No. 7,66-90;

Matches 188; Conservative 28; Mismatches 40; Indels 46; Gaps 6;

QY 2 ASAGGE-----DCSPAPPADR--PHQPFLLGVSGGTASGK 36
Db 56 AGAGGERASVRTSGRGGANHGGRADPAEPHQAQRRRLPYRR-----HGTASGK 109
QY 37 STVCEKIMELLQNEVEQORRVVILSQDRFYKVLTAQKAKALKGQYNPDHPDAPNDL 96
Db 110 SSVCAKIVQLLQNEVEYQKQVILSQDSFYVLTSQKAKALKQFNPDPAPDNEX 169
QY 97 MHTLKNIIVEKTVETPYDVTLSRLBETTVVPAVVLFEGILVYSQEIEMFHLRL 156
Db 170 ILKTLKEITBETKVOIPYDVSLSRKEETVYPAVVLFEGILATYSQEVRLDFQMKL 229
QY 157 FVDTDSVRLSRVLRADV-RRGRLDEQILTYTTFVPAPEFCLPTKKYADVIIPPGVD 215
Db 230 FVDTDSVRLSRVLRADV-RRGRLDEQILTYTTFVPAPEFCLPTKKYADVIIPPGVD 289
QY 216 NMVAINIYVQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESSSR 275
Db 290 NMVAINIYVQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESSSR 335
QY 276 PH 277
Db 336 PH 337

RESULT 3
US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Gluckemann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: US8, THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Query Match 29.6%; Score 428.5; DB 9; Length 125;
Best Local Similarity 67.2%; Pred. No. 56-39;
Matches 84; Conservative 12; Mismatches 28; Indels 1; Gaps 1;

QY 154 LRIFVDTSDVRLSRVLRADV-RRGRLDEQILTYTTFVPAPEFCLPTKKYADVIIPR 212
Db 1 LKIFVDTSDVRLSRVLRADV-RRGRLDEQILTYTTFVPAPEFCLPTKKYADVIIPR 60
QY 213 GVNMMVAINIYVQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESS 272
Db 61 GSDNHVAIDLIVQIIDLINGDICKMRGSGNGRSYKRTFSPGDHPGMLTSGKSHLESS 120
QY 273 SSRPH 277
Db 121 SSRPH 125

RESULT 4
US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Query Match 22.6%; Score 327; DB 12; Length 60;


```
RESULT 8
US-10-156-761-12495
; Sequence 12495, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12495
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12495

Query Match      10.3%; Score 149; DB 15; Length 329;
Best Local Similarity 27.1%; Pred. No. 1.3e-07;
Matches 62; Conservative 36; Mismatches 89; Indels 40; Gaps 11;

QY 23 PELGVGGTASGKSTVCEKIMELGQNEVEQRQRKVVLSQDFYKVLTAEQAKALKG 82
DB 100 PVIYGVAGSVAVGKSTVAFRLQLLSRPEHPR---VELVTTDGL-LPTKELEARGLMS 155
QY 83 QYNFHDHDAFNDMHRTLKNIIVEGK-TVEVPTDPTVTHSLP--ETTYYVPADVLFEG 139
DB 156 RKGF--PESDYRRLTLTFVADIKGKQEBVAPVYSHLYIDVPEQKLTVRP-DILIVEG 212
QY 140 ILVFSQ-----EIRDMFHLRLFVDT-DSVLR--LSR-RVLR----- 172
DB 213 LNVLPALPGKDGRTVGLADYFDFSVYVDARPEDIRRWYLNFRKLRAFAQNPSSYFR 272
QY 173 ----DYRGRDLQTLTQYTFVKAPEFCLPTKKYADVILPRVDMMV 218
DB 273 RYGVSEDAEDALDYARTWTITNKVLLDENAVPTGRALVVRKGPDKV 321

RESULT 9
US-09-734-569-88
; Sequence 88, Application US/09734569
; Patent No. US20020064816A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Citrus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; APPLICANT: Reeki, Ralf
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins involved
; FILE REFERENCE: BASF-NAE-1332-99-US
; CURRENT APPLICATION NUMBER: US/09/734, 569
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/171,101
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentin Ver. 2.1/WordPerfect 6.1
```

```
; SEQ ID NO 88
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-569-88

Query Match      8.2%; Score 119.5; DB 9; Length 143;
Best Local Similarity 26.8%; Pred. No. 6.4e-05;
Matches 37; Conservative 28; Mismatches 52; Indels 21; Gaps 5;

QY 25 LIGVSGTASGKSTVCEKIMELG-----QNEVEQRQRKVVLSQDFYKVL 72
DB 14 VIGLADSGCGKSTFMRRLTSVGCATPCKGNPDNLTISPTTVICL--DDYHSLDR 71
QY 73 AEQKALKGQYNFDPDAFNDMHRTLKNIIVEGKTVFVPTVDPYTH-SRLPETTVVP 131
DB 72 YGRKEKAVNAL----DPRANFDMVEQVKALEKGSVEKPYNVHTGLDAPET--IHP 125
QY 132 ADVLFEGLVFSQDIR 149
DB 126 PKLIVIEGLHPMYDERVR 143

RESULT 10
US-10-301-822-217
; Sequence 217, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamackar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEMO1-029P2RM
; CURRENT APPLICATION NUMBER: US/10/301, 822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339, 971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361, 978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381, 988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 217
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-217

Query Match      6.7%; Score 97; DB 12; Length 216;
Best Local Similarity 24.4%; Pred. No. 0.036;
Matches 53; Conservative 36; Mismatches 64; Indels 64; Gaps 12;

QY 2 ASAGE-DCSPAPEADRHQRFLVGGTASGKSTVCEKIMELGQ-----NEVEQR 55
DB 36 ASAGPERDRAPGAFAL-----GTWGAHSASEVRELBGKTGSSDDIEQL 81
QY 56 ORKVVLSQDFYKVLTAEQAKALKGQYNFDPDA-----FNDMLMHRTL 101
DB 82 HRR-----FKQLSGQPTIR---KENFNNVPELEINPYSKTVRAFPOF--RNL 125
QY 102 KNIVEGKTVFVPTVDPYTH--HSRLPETTVVPADVLEBGLVFSQDIRDMFHLRLV 158
DB 126 RKGPSGLADEINDEFILTIMSYFRPIDTMDR-----EYVELSRKELRPLFHM---Y 175
QY 159 DTDSVRLSRVRLVDVRGRD-LEQLTLQYTFVKA 194
```


Db 176 DSDSDGRITLBYRNVMKMSRSCCETLTSTRSLAPFS 212

RESULT 11
US-10-301-822-220

; Sequence 220, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Berger, Allison

; APPLICANT: Guillemette, Tracy L.

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Thibodeau, Stephen N.

; APPLICANT: Burgart, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MEM01-029P2RNM

; CURRENT APPLICATION NUMBER: US/10/301,822

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/339,971

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978

; PRIOR FILING DATE: 2002-03-05

; PRIOR APPLICATION NUMBER: US 60/381,988

; PRIOR FILING DATE: 2002-05-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 220

; LENGTH: 251

; TYPE: PRT

; ORGANISM: Homo Sapiens

; US-10-301-822-220

Query Match 6.6%; Score 95; DB 12; Length 251;
Best Local Similarity 24.6%; Pred. No. 0.075;

Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAPRDRPHQRPLIGVSGTASGKSTVCCKIMELLGQ-----NEVEOR 55

Db 36 ASAGPERDARPGAPAA-----GTWGAHSASEBEVRELEGKTGSSDQIEOL 81

QY 56 QKRVILISDRFYKVTATAGOKAKALKGQYNFDH-PDA-----FNDLMDHRTL 101

Db 82 HRR-----FKQLSGDQPTIR---KENFNVPDLELNPISRKIVRAFFDN---RNL 125

QY 102 KNIVEGKTVVPTDYVT---HSRLPETTVVPADVVLFGILVFSQEIIRDMFHLRLFV 158

Db 126 RKGPSSGLADEINFEDELITMSYFRPIDTMDR-----EYELSRKXKIRLFLFM---Y 175

QY 159 DTDSDVRLSRRLVDRVGRDLQIILT 185

Db 176 DSDSDGRITLBYRNVM-----VEELLS 197

RESULT 12

US-10-106-698-6128

; Sequence 6128, Application US/10106698

; Publication No. US20030109690A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide

; FILE REFERENCE: PA0005P1

; CURRENT APPLICATION NUMBER: US/10/106,698

; CURRENT FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: PCT/US00/26524

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US 60/157,137

; PRIOR FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6128

; LENGTH: 257

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-106-698-6128

Query Match 6.6%; Score 95; DB 15; Length 257;
Best Local Similarity 24.6%; Pred. No. 0.077;

Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAPRDRPHQRPLIGVSGTASGKSTVCCKIMELLGQ-----NEVEOR 55

Db 26 ASAGPERDARPGAPAA-----GTWGAHSASEBEVRELEGKTGSSDQIEOL 71

QY 56 QKRVILISDRFYKVTATAGOKAKALKGQYNFDH-PDA-----FNDLMDHRTL 101

Db 72 HRR-----FKQLSGDQPTIR---KENFNVPDLELNPISRKIVRAFFDN---RNL 115

QY 102 KNIVEGKTVVPTDYVT---HSRLPETTVVPADVVLFGILVFSQEIIRDMFHLRLFV 158

Db 116 RKGPSSGLADEINFEDELITMSYFRPIDTMDR-----EYELSRKXKIRLFLFM---Y 165

QY 159 DTDSDVRLSRRLVDRVGRDLQIILT 185

Db 166 DSDSDGRITLBYRNVM-----VEELLS 187

RESULT 13

US-10-301-822-215

; Sequence 215, Application US/10301822

; Publication No. US20030148410A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Berger, Allison

; APPLICANT: Guillemette, Tracy L.

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Schlegel, Robert

; APPLICANT: Monahan, John E.

; APPLICANT: Thibodeau, Stephen N.

; APPLICANT: Burgart, Lawrence J.

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; FILE REFERENCE: MEM01-029P2RNM

; CURRENT APPLICATION NUMBER: US/10/301,822

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 60/339,971

; PRIOR FILING DATE: 2001-12-10

; PRIOR APPLICATION NUMBER: US 60/361,978

; PRIOR FILING DATE: 2002-03-05

; PRIOR APPLICATION NUMBER: US 60/381,988

; PRIOR FILING DATE: 2002-05-20

; NUMBER OF SEQ ID NOS: 228

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 215

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Homo Sapiens

; US-10-301-822-215

Query Match 6.6%; Score 95; DB 12; Length 267;
Best Local Similarity 24.6%; Pred. No. 0.082;

Matches 51; Conservative 35; Mismatches 53; Indels 68; Gaps 12;

QY 2 ASAGGE-DCESPAPRDRPHQRPLIGVSGTASGKSTVCCKIMELLGQ-----NEVEOR 55

Db 36 ASAGPERDARPGAPAA-----GTWGAHSASEBEVRELEGKTGSSDQIEOL 81

QY 56 QKRVILISDRFYKVTATAGOKAKALKGQYNFDH-PDA-----FNDLMDHRTL 101

Db 82 HRR-----FKQLSGDQPTIR---KENFNNVDELINPIRSKIYAFEDN---FNL 125
Qy 102 KNIVEGKTEVETVYFVT---HSRLPETTVYVPADVLFEGILVYSOEIRDMFLRLFV 158
Db 126 RKGPGGLADEINFEDEFLTIMSYFRPIDTMD-----EYVESRKELKFLFRM---Y 175
Qy 159 DTSDVRLSRVLRDVRGRDLEQILT 185
Db 176 DSDSDGRITLBEYRNV-----VEELIS 197

RESULT 14

US-09-769-744A-64
Sequence 64, Application US/09769744A
Publication No. US20030134407A1
GENERAL INFORMATION:
APPLICANT: Le Page, Richard WF
APPLICANT: Wells, Jeremy M
APPLICANT: Hambrico, Sean B
APPLICANT: Hambrico, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PMC/P21122MO
CURRENT APPLICATION NUMBER: US/09/769,744A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: PCT/GB99/02452
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: GB 9816336.3
PRIOR FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: US 60/125329
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 196
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 64
LENGTH: 210
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-769-744A-64

Query Match 6.4%; Score 93; DB 12; Length 210;

Best Local Similarity 21.0%; Pred. No. 0.095; Mismatches 79; Indels 66; Gaps 11;
Matches 49; Conservative 39; Mismatches 79; Indels 66; Gaps 11;

Qy 25 LIGVSGTASGKSTVCEKIMELGONEVEORQRKVILISOD-----REFYKVLTA 73
Db 13 IIGITGIGIAGSKSTVYNFL-----RQGGFQVVDADAVVHQLQKPGGRLEFALVQ 61
Qy 74 E--QKAKALKGQY-----FDHPDAFNDLHARTLAKIVGKTV--EYPTVDFVTHS 121
Db 62 HFGQELIENGEILRPLIASLIFSND-----EREWSKOIGELIREELAT---LRE 110
Qy 122 RLPEFTVYVPADVLFEGILVYSOEIRDMFLRLFVDTSDVRLSRVLRDVRGRDLE 181
Db 111 QLAQTEIEFFMDIPL-----LFEDYSDFWFAETWLVYVDDAQVERLMKRDQSKDEAE 164
Qy 182 -QILTQYTFVKPAEEFCLEPTKKYADVIRGVNMMVAINLIVQHIQDILNG 233
Db 165 SRLAQQW-----PLEKKKD-LASQVLDNNGNQQLNQVHILLEG 203

RESULT 15

US-09-815-242-5691
Sequence 5691, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21,078
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5691
LENGTH: 433
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5691

Query Match 6.4%; Score 93; DB 9; Length 433;

Best Local Similarity 23.8%; Pred. No. 0.28; Mismatches 93; Indels 42; Gaps 9;
Matches 53; Conservative 35; Mismatches 93; Indels 42; Gaps 9;

Qy 31 GTASGKSTVCEKIMELGONEVEORQRKVILISQDRFYKVLTAQKAKALKGQYNFDDHP 90
Db 174 GSTHSNSTWCKRSTLAEQLKI---QLEVV-----ESPLHAFETNSSL-----YVHP 218
Qy 91 AFNDLHARTLAKIVGKTEVETVYFVTHSRLEP-TVYVPADVLFEGILVYSOEIR 149
Db 219 LFMNDF---SLKAIPEG--TDVPEVYVYKLFPEEGPITMTILIREWRLMWMKEMAILQAFRV 273
Qy 150 DMFHLRFVDTSDVRLSRVLRDVRGRDLEQILTQYTFVKPA----- 194
Db 274 SVNLQPMWKENYPRRPETIDEGDIEHFILPDILEYLVYKVTAILIDPFSQPDENGH 333
Qy 195 -FEFCLEPTKKY---ADVI-IPRGVNNVAINLIVQHIQDIL 231
Db 334 YFDFSAVPFKQYKNGQDVVQIPRMPSEDYVRRAMIGHIKKML 376

Search completed: November 25, 2003, 08:08:11
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:01:09 ; Search time 20 Seconds
(without alignments)
586.005 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 1450
Sequence: 1 MASAGEDCESPAPADRP...HPGMLTSGKRSHESSRPH 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep: *
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfile01.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1363	94.0	260	4 US-09-536-647-2	Sequence 2, Appl1
2	1335	91.4	277	4 US-09-536-647-3	Sequence 3, Appl1
3	451	31.1	216	4 US-09-134-001C-3618	Sequence 3618, Ap
4	434	29.9	210	4 US-09-107-532A-5024	Sequence 5024, Ap
5	412	28.4	222	4 US-09-198-452A-783	Sequence 783, Ap
6	99.5	6.9	208	4 US-09-134-001C-3785	Sequence 3785, Ap
7	93.5	6.4	216	4 US-09-252-991A-26519	Sequence 26519, A
8	91.5	6.3	1376	1 US-08-420-235B-3	Sequence 3, Appl1
9	91.5	6.3	1376	3 US-08-793-624-3	Sequence 3, Appl1
10	91.5	6.3	1376	5 PCT-US95-10194-3	Sequence 3, Appl1
11	89	6.1	517	4 US-09-252-991A-22249	Sequence 22249, A
12	88	6.1	674	3 US-08-961-083-200	Sequence 200, App
13	88	6.1	674	4 US-09-536-784-200	Sequence 200, App
14	86	5.9	542	4 US-09-198-452A-1003	Sequence 1003, Ap
15	85	5.9	802	4 US-09-134-001C-3840	Sequence 3840, Ap
16	85	5.9	810	4 US-09-540-824-25	Sequence 25, Appl
17	84.5	5.8	647	4 US-09-134-001C-5458	Sequence 5458, Ap
18	84.5	5.8	2639	3 US-09-080-983-3	Sequence 3, Appl1
19	84	5.8	687	2 US-08-540-804-4	Sequence 3, Appl1
20	84	5.8	687	2 US-08-218-265-4	Sequence 4, Appl1
21	84	5.8	687	3 US-08-521-872-4	Sequence 4, Appl1
22	84	5.8	687	3 US-08-590-399-4	Sequence 4, Appl1
23	81	5.6	937	4 US-09-107-532A-4335	Sequence 4335, Ap
24	80.5	5.6	625	4 US-09-252-991A-20865	Sequence 20865, A
25	80.5	5.6	661	4 US-09-328-352-5666	Sequence 5666, Ap
26	80.5	5.6	722	4 US-08-817-832B-32	Sequence 32, Appl
27	80	5.5	205	4 US-09-328-352-8081	Sequence 8081, Ap

28	78.5	5.4	284	4 US-08-858-207A-346	Sequence 346, App
29	78.5	5.4	285	4 US-09-328-352-4464	Sequence 4464, Ap
30	78.5	5.4	414	4 US-09-647-540A-6	Sequence 6, Appl1
31	78.5	5.4	414	4 US-10-119-600-6	Sequence 6, Appl1
32	78.5	5.4	607	3 US-08-781-891-75	Sequence 75, Appl
33	78.5	5.4	607	3 US-09-618-166-75	Sequence 75, Appl
34	78.5	5.4	722	4 US-09-984-890-4	Sequence 4, Appl1
35	78.5	5.4	724	4 US-09-984-890-2	Sequence 2, Appl1
36	78.5	5.4	745	4 US-09-523-843-36	Sequence 36, Appl
37	78.5	5.4	779	1 US-08-426-627-4	Sequence 4, Appl1
38	78.5	5.4	779	1 US-08-426-627-24	Sequence 24, Appl1
39	78.5	5.4	836	1 US-08-426-627-6	Sequence 6, Appl1
40	78.5	5.4	837	1 US-08-426-627-23	Sequence 23, Appl1
41	78.5	5.4	2482	1 US-08-328-254-6	Sequence 6, Appl1
42	78.5	5.4	3248	1 US-08-353-700-1	Sequence 1, Appl1
43	78.5	5.4	3248	5 PCT-US95-16216-1	Sequence 1, Appl1
44	78	5.4	237	3 US-09-320-878-11	Sequence 11, Appl
45	78	5.4	237	3 US-09-105-537-18	Sequence 18, Appl

ALIGNMENTS

```
RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-2

Query Match          94.0%; Score 1363; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 1e-150;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGEDCESPAPADRP...PFLIGVSGTASGKSTVCCKIMELLGQNEVEQRGRKV 60
   |||
Db 1 MASAGEDCESPAPADRP...PFLIGVSGTASGKSTVCCKIMELLGQNEVEQRGRKV 60
   |||

QY 61 IISDRFYKVLTAEOAKALKGQYNFDPDADNDIMRTLKQIVGKTVVPTYDFVTH 120
   |||
Db 61 IISDRFYKVLTAEOAKALKGQYNFDPDADNDIMRTLKQIVGKTVVPTYDFVTH 120
   |||

QY 121 SRLPETTVYPPADVLFESGILVFYSOETRDMDHLRLFVDTSDVRSRVLDVVRGRDL 180
   |||
Db 121 SRLPETTVYPPADVLFESGILVFYSOETRDMDHLRLFVDTSDVRSRVLDVVRGRDL 180
   |||

QY 181 EQLITQYTFVAPAFEEFCLPTKRYADVILIPRGVDMVAINILVQHIQDILNGDICKMR 240
   |||
Db 181 EQLITQYTFVAPAFEEFCLPTKRYADVILIPRGVDMVAINILVQHIQDILNGDICKMR 240
   |||

QY 241 GGSNGRSYKRTSPSGDHPG 260
   |||
Db 241 GGSNGRSYKRTSPSGDHPG 260
   |||

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
```

APPLICANT: Johnson, Randell
TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
FILE REFERENCE: GP50020
CURRENT APPLICATION NUMBER: US/09/536,647
CURRENT FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 277
TYPE: PRT
ORGANISM: Human
US-09-536-647-3

Query Match 91.4%; Score 1325; DB 4; Length 277;
Best Local Similarity 91.7%; Pred. No. 3e-146;
Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

QY 1 MASAGDECSAPAEADRPORPPLIGVSGTASGKTCVKIMELLGONEVEORORRV 60
DB 1 MASAGGGSSEAPAEADRPORPPLIGVSGTASGKTCVKIMELLGONEVEORORRV 60
QY 61 ILSQDREYKVLTAQOKAKALKQGYNPDHDPDAFNDMLMRTLKNIVEGKTVETPTDEVT 120
DB 61 ILSQDCFYKVLTAQOKAKALKQGYNPDHDPDAFNDMLMRTLKNIVEGKTVETPTDEVT 120
QY 121 SRLPETHVVPADVVLREGILVFESQERDMFHLRLPVDITSDVRLSRVLRDVRGRDL 180
DB 121 SRLPETHVVPADVVLREGILVFESQERDMFHLRLPVDITSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTTFAVKPAFEFCLPTKKYADVIIPRGVNNVAINLYQHIQDILNGDCKMR 240
DB 181 EQLITQYTTFAVKPAFEFCLPTKKYADVIIPRGVNNVAINLYQHIQDILNGDCKMR 240
QY 241 GGSNGRSYKRTSEPGDHGMLTSGKSHLESSSRPH 277
DB 241 GGPNGRNHKKRTFPEPGDHPGLATGKSHLESSSRPH 277

RESULT 3
US-09-134-001C-3618
Sequence 3618, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3618
LENGTH: 216
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3618

Query Match 31.1%; Score 451; DB 4; Length 216;
Best Local Similarity 42.9%; Pred. No. 2.4e-44;
Matches 91; Conservative 44; Mismatches 61; Indels 16; Gaps 4;
QY 25 LIGVSGGTASGKTCVKIMELLGONEVEORORRVILSODRFYK--VITAEQAKALK 81
DB 15 IIGIAGSGSGKTVTAIMKNL-----EGHSVALLAODYKKDSHLSFEERLET-- 65
QY 82 GQVYFDPADFDNDMLMRTLKNIVEGKTVETPTDFTVTHSRLPETHVVPADVVLREGIL 141
DB 66 ---NYDHPFAFDNDLLIHNLDLRNGKPEVETPDYSQHSKSTETIFDPPDVIVIGIF 122
QY 142 VFVYSQERDMFHLRLPVDITSDVRLSRVLRDVR-RGRDLQILITQYTTTFVKPAFEFCL 200

DB 123 ALENNTLRDMMDVKIYVDDADARILRLTRDKEGRMTSEVINGYLVNVRMHEQFIE 182
QY 201 PTKKADVIIIPRGVNNVAINLYQHIQDILN 232
DB 183 PTKKHADIIIPGGSNKVADIMWTTKIQSLVS 214

RESULT 4
US-09-107-532A-5024
Sequence 5024, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 5024:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...210
SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
US-09-107-532A-5024

Query Match 29.9%; Score 434; DB 4; Length 210;
Best Local Similarity 38.6%; Pred. No. 2.2e-42;
Matches 83; Conservative 51; Mismatches 65; Indels 16; Gaps 4;
QY 22 RPLIGVSGGTASGKTCVKIMELLGONEVEORORRVILSODRFYK--VITAEQAKALK 78
DB 6 KPIIGVGTGSGSGKTSVRAIFNPPDH-----SIMMLEQDSYKKDSHLSFEERL 58
QY 79 ALKGQYNPDHDPDAFNDMLMRTLKNIVEGKTVETPTDFTVTHSRLPETHVVPADVVLREG 138
DB 59 T-----NYDHPFAFDNDLLIHNLDLRNGKPEVETPDYSQHSKSTETIIOEKEVILIE 113
QY 139 GILVFYSQERDMFHLRLPVDITSDVRLSRVLRDVR-RGRDLQILITQYTTTFVKPAFEF 197

Db 114 GILLLEDERLDMJKNYVDDDDIRIRKRMERKGLBSVLEQYLTVCWYHQ 173

Qy 198 FCLPTKKYADVILPRGVDMMAINLIVQHIDILN 232
| |||:|:|:|:|:|:|:|:|:|
Db 174 FIETPKRYADIIYPBEGGHNVAIDLTTVASFLN 208

```

RESULT 5
US-09-198-452A-783
Sequence 783, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 783
LENGTH: 222
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-783

```

Query Match	28.4%	Score 412	DB 4	Length 222
Best Local Similarity	38.4%	Pred.No. 9.1e-40		
Matches 84	Conservative 54	Mismatches 59	Indels 22	Gaps 6

```

QY 25 LIAGSGTAGKSTVECKIMELLGONEVBOROKVILISODEFY---VLTAEOCKAKK 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 8 IIGTGGSGAGKTTLQNIKEIFGD-----SVIQDNYVYDRSHRYPEEAGNI I- 58

QY 82 GQNFHPDAFENDMLHRLKNI VEKTYEVPYDPYTHSR-LPETVYVPADVLFEFGI 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 ----WDHPDAFENDMLISDKIKLNNEIYQAFEDFVLNNSKSTLETTYSYVLVGGI 11
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Qy 141 LVFVSGEIDMFHLRLFVDTSDSVRLSRVLADV -RGRLDQLITQYTTFKKPAFEFC 199

Db 115 LVFENQELRLMDIRIFVDTYDADERILRWVRDVQEGGSDVDCIMSRYLSTWYPMHEKFI 174

```
QY      200 LPTKKYADVII PRGVDNVAI NLIVQ---HIQDILNGD 234
          ||:|||::: ||::|
Db      175 EPTRYADIIVHGNYRQNVNTNLSQKIKHLENALES D 213
```

```

RESULT 6
US-09-134-001C-3785
; Sequence 3785, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucetle-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3785
; LENGTH: 208
; TYPE: PR1
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3785

```

Query Match	6.9%	Score	99.5	DB	4	Length	208
Best Local Similarity	21.5%	Pred. No.	0.0026				
Matches	40	Conservative	48	Mismatches	71	Indels	27
						Gaps	8

```

OY 25 :IGVSGGTAASKRSTYCE-----KIMEL-LGQNEVEQRQRVYLLSQDRFVKVLTAE-- 74
Db 9 VIGITGGTATGTSVSELTITAYGFKEIVADIASRAVKKGSKGLEQVEKI FEEBALDENG 68
OY 75 QKAVALKQGVYFHDHDAFD--NDLMHRTLKNIVEGK-----TVEVP-TYDFVTH 120
Db 69 EMNNQYVGEIYFNNHDLREALNEIHPYVREIMBEQKNVLEHGHHVIMDIPLYTENELQ 128
OY 121 SRLPETTVYVPADVVLTFGGILVFVSQELRDMFHLRF--VDTDSVRLSRRLVDRVRRR 178
Db 129 DTVEVEVVVVVTSYSIQIDRLMERNNLSLEDA-KARVYSQISIDDKSRMADHVIDNLGDKYL 187
OY 179 DLEQIL 184
Db 188 ELKQWL 193

```

```

RESULT 7
US-09-252-991A-26519
: Sequence 26519, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIORITY FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIORITY FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIORITY FILING DATE: 1998-07-27
: NUMBER OF SEQ. ID NOS: 33142
: SEQ. ID NO 26519
: LENGTH: 216
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26519

```

Query Match	6.4%	Score 93.5	DB 4	Length 216
Best Local Similarity	23.7%	Pred. No. 0.014		
Matches 41	Conservative 31	Mismatches 74	Indels 27	Gaps 6

```

Qy 22 RPLFISGSGTISGKSTVEKIMEL-LGQNEVEGQQRKVVLSQRFKV-----L 71
Db 16 QPWLILGLTGISGSKAAAEHPISLGVHLVDAADHARVWVEGBRGALAKIYERFGDILL 75
Qy 72 TAPQAKALKGQYNFEDHPDAFNDLHRTIKNIYEGKTEVEPTYDFVTHSLRPETTVYR 133
Db 76 PDGQDLRALRREIFAPR-ERRWLEQLNHLGLGAEIVQ-----YLARASSPAIIVSP 128
Qy 132 ADVVLEFEGILVFYSQEIRDMFLRLYFVPTDSDVLRSSREVLRLDVRGRDLEQIL 184
Db 129 -----LVESGQ--RQMTIRVLVVDIPHLIQRLQRTMLDKYSSQEVRSIL 171

```

RESULT 8
US-08-420-235B-3
Sequence 3, Application US/08420235B
Patent No. 5801042
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

Qy	32	TAGC-----	-SWCEKIMELLIGONEVEOROKVILISQDFYKULTAEOKAKLKG	82
		:	:	:
Db	266	TAGCGQVAGVLETTDSVVRKLNILLGQESAMSGPAVAASVVRGAMLVTAVSYGRAMRN	325	
Qy	83	QYNF-----	DHPAEPNDLMHRTLKNIYEG-KTYEVPFYDFVTHSRLEPFTYVYPAVVL	136
		:	:	:
Db	326	FEQFMARIYDHPALPS-----	-VEGDPAALADGHDIEIORTKTAASLVKIGDKFVA	373
Qy	137	FEGLIFVYSQEIOWMFMHRLFVYDTDSDVRLSRVRLDVRRGRDLEQILTYOTTFVKEAFE	166	
Db	376	IESL-----	-GEMYNETOFCPCPLNRI-----	407
Qy	197	EFLPLTKKYADVILPRGVDNMVAIMLYIQHODILNDIC-----	KMHVGS	243
		:	:	:
Db	408	--LPVPRISTSVSRVGS-PALQSTETVYVKNVFLCEGYONALKSICHPMHNPQ	463	
Qy	244	NGRSYKRTSEP-GDH	258	
		:	:	:
Db	464	SAQALNQAFPPDPGGH	479	
RESULT 10				
PCT-US95-10194-3				

1 GENERAL INFORMATION:
2 APPLICANT: The Trustees of Columbia University in the City of New York
3 APPLICANT: City
4 TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
5 TITLE OF INVENTION: SEQUENCES AND USES THEREOF
6 NUMBER OF SEQUENCES: 45
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Cooper & Dunham LLP
9 STREET: 1185 Avenue of the Americas
10 CITY: New York
11 STATE: New York
12 COUNTRY: U.S.A.
13 ZIP: 10036
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: PatentIn Release #1.25
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: PCT/US95/10194
21 FILING DATE:
22 CLASSIFICATION:
23 ATTORNEY/AGENT INFORMATION:

```

REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPM/MS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-10194-3

Query Match      6.3%; Score 91.5; DB 5; Length 1376;
Best Local Similarity 20.7%; Pred. NO. 0.44;
Matches 53; Conservative 36; Mismatches 96; Indels 71; Gaps 10

Oy      32 TASGK-----STVCEKIMELLIGQNEVROQRKVILTSQDFYKLTVEQKAKLKG 82
      |||||:      :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      266 TAAQQQVAGVLETTDSVVRKRLNMLLGGQVESAMSGPAALAASTVVGANLVTVASVGRAMRN 325
      |||||:      :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy      83 QYVF-----DHPDAFNDIMLRFTLKNIYEG-KTYEVEVPYIDTVTHSRLPETTYVVPADVL 136
      |||||:      :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      326 FEQFMARIVDHPNAPLS-----VEGDKPALADGDHDIQRTITLAASLVKIGDKFEVA 375
      |||||:      :|||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
QY 137 FEGILVFYSQIRDMFHLRLFDVTDSDVRLSRVLRDVRGRDLQILTYTTFVKPAFE 196
| : : : : :
DB 376 IESL-----QKRYNERQFPCPLNRR-----QYTFPFVGLH 407
QY 197 EFCGLPTKRYADVIIIRGVDNMVAIINLIYOHIODLNGDIC-----KMHKGS 243
| : : : : :
DB 408 ---LPIVRYSTSVSRVES-PAIQSTETVYVNNKNVPLFCGYONALKSIChPRMHNPTQ 463
QY 244 NGRSYKRTFSEP-GDH 258
| : : : : :
DB 464 SAQALNQAFPPDQGH 479

RESULT 11
US-09-252-991A-22249
; Sequence 22249, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22249
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22249

Query Match
Best Local Similarity 23.3%; Pred. No. 0.18;
Matches 59; Conservative 30; Mismatches 102; Indels 62; Gaps 11;

QY 32 TASGKSTVCEKIMELGONEVEQR-----QKRVILSQDRPFYKVLNAEQAKALKGOYN- 85
| : : : : :
DB 56 TAAISQALKRKRYIKERIPSEAEKRONFHYPLRICYRCQGEVPTTRAYAKVQ-VPGESYV 114
QY 86 -FPHPDFAFDNDLHRTKNTVEGKTVVPTVDFTYTRRLBETTYVYVADVVLPFGILV-- 142
| : : : : :
DB 115 TWTHPDAFRKYLLEQ-LRPLMDFTYVQVGP--SQODIPRYVVEQDELAGSGVYAAE 171
QY 143 -----FYSGEIRDMFHLRLFDVTDSDVRLSRVLRDVRGRDLQ 182
DB 172 LARVFPSTDLASASDDIADGLYERADQLPLALFDAARVDFSLRLVHYTGSWMKRVQ 231
QY 183 -ILTYOTTFV-----KPAEFECLEPTKRYADVIIIRGVDNMVAIINLIYOH 227
| : : : : :
DB 232 WILTYNHRVYDQFIRLGLTRLEDRPFVWMLP-----GNVILIRGMDGEAA-NAIVAGV 286
QY 228 QDILNEDICMHR 240
| : : : : :
DB 287 -----VWHR 290

RESULT 12
US-08-961-083-200
; Sequence 200, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
```

```
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 200:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 674 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-200

Query Match
Best Local Similarity 27.5%; Pred. No. 0.36; Length 674;
Matches 30; Conservative 16; Mismatches 35; Indels 28; Gaps 6;

QY 105 VEGKTVVPTYDF-----VTHSRLEPTTYVYVADVLEFGLVIFYSGEIRDMFHLRLPV 158
| : : : : :
DB 562 IETLVSLDPIQKIAKARIIVSDRLAQMT-IYD-NVTIFNGI-----HYLV 605
QY 159 DTDSVRLSRVLRDVRGRDLQILTYTTFVKP--ABEFCLEPTKRY 205
| : : : : :
DB 606 DVNDELVETSOVLIDINHGKTEBILDOFANLGRKPLTSFEN---TKTY 650

RESULT 13
US-09-536-784-200
; Sequence 200, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: November 25, 2003, 07:54:09 ; Search time 40 Seconds
(without alignments)
1099.182 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGGEDPESPAEADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.GeneSeq.19Jun03:*

- 1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1450	100.0	277	22	AA19502 Human polypeptide
2	1450	100.0	277	22	AA19502 Human polypeptide
3	1450	100.0	277	22	AA19502 Human polypeptide
4	1450	100.0	277	22	AA19502 Human polypeptide
5	1372.5	94.7	276	22	AA19502 Human polypeptide
6	1363	94.0	260	22	AA19502 Human polypeptide
7	1325	91.4	277	22	AA19502 Human polypeptide
8	941	64.9	190	23	AA19502 Human polypeptide
9	934	64.4	261	22	AA19502 Human polypeptide

10	934	64.4	261	22	AA19502 Human polypeptide
11	901	62.1	337	23	AA19502 Human polypeptide
12	886	61.8	337	22	AA19502 Human polypeptide
13	889	61.3	335	22	AA19502 Human polypeptide
14	738	50.9	326	22	AA19502 Human polypeptide
15	589	40.6	326	22	AA19502 Human polypeptide
16	579	39.9	120	23	AA19502 Human polypeptide
17	578	39.9	481	21	AA19502 Human polypeptide
18	578	39.9	490	21	AA19502 Human polypeptide
19	578	39.9	512	21	AA19502 Human polypeptide
20	570.5	39.3	500	21	AA19502 Human polypeptide
21	570.5	39.3	511	21	AA19502 Human polypeptide
22	545.5	37.6	548	23	AA19502 Human polypeptide
23	542	37.4	614	22	AA19502 Human polypeptide
24	542	37.4	614	22	AA19502 Human polypeptide
25	540.5	37.3	548	22	AA19502 Human polypeptide
26	540.5	37.3	578	22	AA19502 Human polypeptide
27	529	36.5	415	21	AA19502 Human polypeptide
28	529	36.5	433	21	AA19502 Human polypeptide
29	529	36.5	441	21	AA19502 Human polypeptide
30	529	36.5	443	21	AA19502 Human polypeptide
31	529	36.5	461	21	AA19502 Human polypeptide
32	529	36.5	469	21	AA19502 Human polypeptide
33	526.5	36.3	540	21	AA19502 Human polypeptide
34	518	35.7	477	21	AA19502 Human polypeptide
35	518	35.7	500	21	AA19502 Human polypeptide
36	517.5	35.7	466	21	AA19502 Human polypeptide
37	507.5	35.0	476	21	AA19502 Human polypeptide
38	507.5	35.0	498	21	AA19502 Human polypeptide
39	507.5	35.0	499	21	AA19502 Human polypeptide
40	507.5	35.0	521	21	AA19502 Human polypeptide
41	507	35.0	465	21	AA19502 Human polypeptide
42	507	35.0	487	21	AA19502 Human polypeptide
43	490.5	33.8	274	21	AA19502 Human polypeptide
44	490.5	33.8	292	21	AA19502 Human polypeptide
45	490.5	33.8	300	21	AA19502 Human polypeptide

ALIGNMENTS

RESULT 1	AA19502	standard; Protein; 277 AA.
ID	AA19502	
AC	AA19502	
XX	22-OCT-2001	(first entry)
DE	Human polypeptide SEQ ID NO 2647.	
XX	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocarcin;	
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia.	
OS	Homo sapiens.	
XX	WO200153312-A1.	
PN	26-UTL-2001.	
PD	26-DEC-2000; 2000US-054263.	
XX	21-JAN-2000; 2000US-0488725.	
XX	25-APR-2000; 2000US-0552317.	
PR	09-UTL-2000; 2000US-0598042.	
PR	19-UTL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	

PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI58658.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX Example 4; SEQ ID NO 2647; 10078pp; English.
 PS
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX Sequence 277 AA;
 SQ
 Query Match 100.0%; Score 1450; DB 22; Length 277;
 Best Local Similarity 100.0%; Pred. No. 3.2e-132;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 DB 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 QY 61 ILSQDRFYKVLTAQKAKALKGQYNFPHDAFNDMLMRTLKNIVEGTVVPTVDFVTH 120
 DB 61 ILSQDRFYKVLTAQKAKALKGQYNFPHDAFNDMLMRTLKNIVEGTVVPTVDFVTH 120
 QY 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDMVAINALIVQHIODILNGDICKMHR 240
 DB 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDMVAINALIVQHIODILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
 DB 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
 RESULT 2
 AAB93941
 ID AAB93941 standard; Protein; 277 AA.
 XX
 AC AAB93941;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:13952.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.

XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 13952; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH1632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 XX Sequence 277 AA;
 SQ
 Query Match 100.0%; Score 1450; DB 22; Length 277;
 Best Local Similarity 100.0%; Pred. No. 3.2e-132;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 DB 1 MASAGGEDCSPAPADRPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORRKYV 60
 QY 61 ILSQDRFYKVLTAQKAKALKGQYNFPHDAFNDMLMRTLKNIVEGTVVPTVDFVTH 120
 DB 61 ILSQDRFYKVLTAQKAKALKGQYNFPHDAFNDMLMRTLKNIVEGTVVPTVDFVTH 120
 QY 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVYPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDMVAINALIVQHIODILNGDICKMHR 240
 DB 181 EQILTYTTTVPKPAFEFCFPTKKYADVILIPRGVDMVAINALIVQHIODILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277

DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6219.
XX
XX Human; noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0683036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Aunndi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60444.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6219; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX
SQ Sequence 296 AA;
Query Match 100.0%; Score 1450; DB 22; Length 296;
Best Local Similarity 100.0%; Pred. No. 3.5e-132;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGDECPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 60
DB 20 MASAGDECPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 79
QY 61 ILSODRFYKVLTAEOAKALKGOYNPDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 120
DB 80 ILSODRFYKVLTAEOAKALKGOYNPDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 139

QY 121 SRLPETTVYPADVLFEGILVFYSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 180
DB 140 SRLPETTVYPADVLFEGILVFYSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 199
QY 181 EQILTYTTFVKRPAFEFCPLTKKYADVILPRGVDMVAIINLVQHIOILNGDICKMR 240
DB 200 EQILTYTTFVKRPAFEFCPLTKKYADVILPRGVDMVAIINLVQHIOILNGDICKMR 259
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKRSHLESSSRH 277
DB 260 GGSNGRSYKRTFSEPGDHPGMLTSGKRSHLESSSRH 296
RESULT 5
ID AAG64506 standard; Protein; 276 AA.
XX
XX AAG64506;
AC
XX 02-OCT-2001 (first entry)
DT
XX Human uridine kinase.
DE
XX Human; uridine kinase; UK.
OS
XX Homo sapiens.
XX
PN CN1287172-A.
PD 14-MAR-2001.
XX
PF 07-SEP-1999; 99CN-0118818.
PR 07-SEP-1999; 99CN-0118818.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Zhao Y, Zhang H;
XX
XX WPI; 2001-409529/44.
DR N-PSDB; AAH75355.
XX
PT Human uridine kinase and its coding sequence, preparation and
PT application -
XX
XX Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.
XX
XX The invention relates to human uridine kinase (UK).
XX
SQ Sequence 276 AA;
Query Match 94.7%; Score 1372.5; DB 22; Length 276;
Best Local Similarity 96.0%; Pred. No. 1e-124;
Matches 266; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
QY 1 MASAGDECPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 60
DB 1 MASAGDECPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEORQRKV 59
QY 61 ILSODRFYKVLTAEOAKALKGOYNPDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 120
DB 60 ILSODRFYKVLTAEOAKALKGOYNPDHPDAFNDMLMRTLKNIIVEGKTVVPTDYFVTH 119
QY 121 SRLPETTVYPADVLFEGILVFYSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 180
DB 120 SRLPETTVYPADVLFEGILVFYSOEIRDMFRLRLFVTDSDVRLSRVLRDVRGRD 179
QY 181 EQILTYTTFVKRPAFEFCPLTKKYADVILPRGVDMVAIINLVQHIOILNGDICKMR 240
DB 180 EQILTYTTFVKRPAFEFCPLTKKYADVILPRGVDMVAIINLVQHIOILNGDICKMR 239
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKRSHLESSSRH 277

CC immunological disorders. This sequence represents the protein relating to
 CC a human uridine kinase (UDK) of the invention.
 CC NOTE: The present sequence is stated as being the same as that shown as
 CC SEQ ID NO. 2 in the sequence listing of the specification. However, the
 CC sequences differ.
 XX
 SQ Sequence 277 AA;
 Query Match 91.4%; Score 1325; DB 22; Length 277;
 Best Local Similarity 91.7%; Pred. No. 4.3e-126;
 Matches 254; Conservative 11; Mismatches 12; Indels 0; Gaps 0;
 QY 1 MASAGBDCSPAPADRPHPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORQKVV 60
 DB 1 MASAGGGSSESAPADRPQPRPFLIGVSGGTASGKSTVCCKIMELLGONEVDRQRKVV 60
 QY 61 ILSODRFKVLTAEOKAKALKGQYNFDPDAFNDMDMHTLKNIVEGTVVEPTYPVTH 120
 DB 61 ILSODCFYKVLTAEOKAKALKGQYNFDPDAFNDMDMHTLKNIVEGTVVEPTYPVTH 120
 QY 121 SRLPETTVVPADVVLFEGLIVFYSGEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVPADVVLFEGLIVFYSGEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTOYTTFVKPAFEBCLEPTKKYADYIIPRGYDNVAINLIVQHIQDILNGDICWHR 240
 DB 181 EQILTOYTAFAVKPAFEBCLEPTKKYADYIIPRGYDNVAINLIVQHIQDILNGDLCKRHR 240
 QY 241 GGSNGRSKRTFSESPGDHPGMLTSGKSHLESSSRPH 277
 DB 241 GGPNGRNHKKRTFPEPGDHGVLATGKRSKSHLESSSRPH 277
 RESULT 8
 ABB89353
 ID ABB89353 standard; Protein, 190 AA.
 XX
 AC ABB89353;
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1729.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-20551SP.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI: 2002-122016/16.
 DR N-PSDB: ABL89762.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 11; SEQ ID NO 1729; 2081ip + Sequence Listing, English.
 XX

CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 190 AA;
 Query Match 64.9%; Score 941; DB 23; Length 190;
 Best Local Similarity 98.9%; Pred. No. 4.6e-83;
 Matches 183; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MASAGBDCSPAPADRPHPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORQKVV 60
 DB 1 MASAGBDCSPAPADRPHPORPFLIGVSGGTASGKSTVCCKIMELLGONEVEQORQKVV 60
 QY 61 ILSODRFKVLTAEOKAKALKGQYNFDPDAFNDMDMHTLKNIVEGTVVEPTYPVTH 120
 DB 61 ILSODRFYKVLTAEOKAKALKGQYNFDPDAFNDMDMHTLKNIVEGTVVEPTYPVTH 120
 QY 121 SRLPETTVVPADVVLFEGLIVFYSGEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 DB 121 SRLPETTVVPADVVLFEGLIVFYSGEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILT 185
 DB 181 EQILT 185
 RESULT 9
 AAM38694
 ID AAM38694 standard; Protein, 261 AA.
 XX
 AC AAM38694;
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1839.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 XX

PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT,
XX WPI; 2001-442253/47.
DR N-PSDB; AA157850.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 3; SEQ ID NO 1839; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA158642-AA162213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
SQ Sequence 261 AA;
Query Match 64.4%; Score 934; DB 22; Length 261;
Best Local Similarity 69.5%; Pred. No. 3,4e-82;
Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;
QY 8 DCESPAPADRPH-QRPFLIGVSGGTASGKSTVCEKIMELLGQNEVEBQRQKVVILSQDR 66
DB 4 DSEQTQNHQNGGEPFLIGVSGGTASGKSSVCARKIVQLLGQNEVDYRQKVILSQDS 63
QY 67 FYKVLTRPQAKALKGQYNDHPDAPFNDLMHRTLKIVGKYVEVTTYPTVTHSRLEPT 126
DB 64 FYEVLSEQAKALKGQFPNDHPAPFNEMLIKTLKEITGKTVQIPVYFVSHSRKEET 123
QY 127 TVVYPPADVLFEGILVYSGEIRDMFHLRFVPTDSPVRLSRPLRPDV-RRGRDLRQILT 185
DB 124 VVYPPADVLFEGILAFYSQEVKDLFQMKLFPVDIDATRLSRKVLNDISERGDLEQILS 183
QY 186 QYTFVPAPAEFECFLPTKAYADVIIPRGVDNMVAIINIVOHIDILNGDICKMHGRGSG 245
DB 184 QYITFVPAPAEFECFLPTKAYADVIIPRGADNLVAIINIVOHIDILNG-----G 232
QY 246 RSTKRTSEBPDHPCMLTSGKSHLESSSRPH 277
DB 233 PSKRQT---NGCINGYTPSRKRQASESSSRPH 261
RESULT 10
AAB73494
ID AAB73494 standard; Protein; 261 AA.
XX
XX AAB73494;
XX
XX 31-JUL-2001 (first entry)
XX
XX Human transferrase HTFS-1, SEQ ID NO:1.
XX
XX Human transferrase; HTFS; agonist; antagonist; cellular signalling;
KM proliferation; cell proliferative disorder; immune disorder;

KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KM gene therapy; drug screening.
XX
OS Homo sapiens.
XX
XX WO200132888-A2.
PN
PD 10-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-US30485.
PF
XX
XX 04-NOV-1999; 99US-0163595.
PR
XX
XX (INCYTE GENOMICS INC.
PA
XX
XX Tang YT, Yue H, Hillman UL, Lal P, Bandman O, Patterson C;
PI Shih LL, Azimzai Y, Lu DAM, Baughn WR,
XX WPI; 2001-328796/34.
DR N-PSDB; AAH23801.
XX
XX
XX Human transferrase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX
XX Claim 1; Page 103-104; 157pp; English.
PS
XX
XX Sequences AAB73494-AAB73335 represent novel human transferrase proteins
CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HTFS proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferrase activity, or for compounds that specifically bind to an HTFS
CC protein or which modulates the activity of an HTFS protein.
CC Pharmaceutical compositions comprising an HTFS protein, HTFS
CC agonist or antagonist, or genetic construct encoding an HTFS
CC protein are useful for treating a disease or condition associated
CC with decreased or increased expression of functional HTFS. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HTFS may be
CC used for the diagnosis of disorders associated with the expression of
CC HTFS, or in assays to monitor patients being treated with HTFS or
CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.
XX
SQ Sequence 261 AA;
Query Match 64.4%; Score 934; DB 22; Length 261;
Best Local Similarity 69.5%; Pred. No. 3,4e-82;
Matches 189; Conservative 27; Mismatches 40; Indels 16; Gaps 4;
QY 8 DCESPAPADRPH-QRPFLIGVSGGTASGKSTVCEKIMELLGQNEVEBQRQKVVILSQDR 66
DB 4 DSEQTQNHQNGGEPFLIGVSGGTASGKSSVCARKIVQLLGQNEVDYRQKVILSQDS 63

Qy	67	FKVLTAEQKAAKALQGVNFDPHPDADNDNLMTLNIVEGKVEVPTDPVTHSRLET	126
Dd	64	FVRLTSEQKAAKALQGVNFDPHPDADNDNLMTLNIVEGKVEVPTDPVTHSRKEET	123
Qy	127	TVVYPADVVLFEGILAFVSGEIRDMENHLRFPVDTSDVLRSRRLDV--RRGDLBOILL	185
Dd	124	TVVYPADVVLFEGILAFVSGEVRDLFQMKLFVDTADTRLSRRVLNDISRGDLFQILS	183
Qy	186	QYTTTVKRAFEFECFLPTKRYADVITPRGVDNMVAINLVIQHIODILNGDICMHRGSGS	245
Dd	184	QYTTTVKRAFEFECFLPTKRYADVITPRGVDNLVAINLVIQHIODILNG-----G	232
Qy	246	RSYKRTFSEPGDHPGMLTSGKSHLESSSRPH	277
Dd	233	PSKRQT--NGCLNGYTPSRKQAQSESSSRPH	261
RESULT 11			
ID	ABP41393	standard; Protein; 337 AA.	
XX	ABP41393;		
XX	22-AUG-2002	(first entry)	
XX	DE	Human ovarian antigen HOELP29, SEQ ID NO:2525.	
KW	Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;		
KW	ovarian cancer; breast cancer; tumour; reproductive system disorder;		
KW	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;		
KW	PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;		
KW	inflammatory condition; immune disorder; blood disorder;		
KW	cardiovascular disorder; respiratory disorder; neurological disorder;		
KW	gastrointestinal disorder; urinary system disorder; drug screening;		
KW	gene therapy; chromosome mapping; forensic analysis;		
KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;		
KW	antiinflammatory; gynaecological; reproductive; chromosome 1p32.		
OS	Homo sapiens.		
XX	WO200200677-A1.		
XX	03-JAN-2002.		
XX	07-JUN-2001; 2001WO-US18569.		
XX	07-JUN-2000; 2000US-209467P.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PI	Birse CE, Rosen CA;		
XX	WPI; 2002-147878/19.		
XX	N-PSDB; ABO54470.		
PT	Isolated nucleic acid molecules encoding novel ovarian polypeptides,		
PT	useful in the prevention, treatment and diagnosis of cancer (e.g.		
PT	ovarian cancer), immune disorders, cardiovascular disorders and		
PT	neurological diseases -		
PS	Claim 11; SEQ ID NO 2525; 2922pp; English.		
XX	ABP41054-		
CC	ABP413288) and to cDNAs encoding them (ABO54131-ABO56305), and also		
CC	encompasses polypeptides 90% identical and polynucleotides 95% identical		
CC	to the sequences of the invention. The invention additionally relates to		
CC	recombinant vectors and host cells comprising human ovarian antigen		
CC	polynucleotides, antibodies against human ovarian antigens, and the use		
CC	of ovarian antigen polynucleotides and polypeptides in diagnosing,		
CC	treating, prognosing or preventing various ovary and/or breast-related		
CC	disorders. Such conditions include ovarian cancer and breast cancer, and		
CC	metastatic tumours of ovarian or breast origin, reproductive system		

CC		disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC		polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC		disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC		shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC		vaginitis), immune disorders (e.g., congenital and acquired
CC		immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC		blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC		respiratory disorders, neurological disorders, gastrointestinal disorders
CC		and urinary system disorders. Ovarian antigen polypeptides and
CC		polynucleotides may also be used in screening for compounds which
CC		modulate ovarian antigen expression or activity. The polynucleotides may
CC		further be used for gene therapy, chromosome mapping, in the
CC		identification of individuals and in forensic analysis, and the
CC		polypeptides may be used as food additives or to prepare antibodies
CC		useful in disease diagnosis, drug targeting and phenotyping. The present
CC		sequence represents a human ovarian antigen of the invention.
CC		Note: The sequence data for this patent did not form part of the printed
CC		spectification, but was obtained in electronic format directly from WIPO
CC		at ftp.wipo.int/pub/published_pct_sequences.
xx	Sequence	337 AA;
SQ		
Query Match	62.1%	Score 901; DB 23; Length 337;
Best Local Similarity	62.6%	Pred. No. 7.6e-79;
Matches 189;	Conservative 28;	Mismatches 39; Indels 46; Gaps 6;
OY	2 ASAGGE-----DDESPAPEADR--PHORPFLIGVSGTAGSK 36	: :
DB	56 AGAGGERASVTGSGRGCAHNGCGQRADPAEPFAACGRALPYRR-----HGCTAGSK 109	: :
OY	37 STVCEKIMELLIGONEVEBORQKKVILISODRFYKVLTAEOKAKALKGYENPDHPDAFENDL 96	: :
DB	110 SSVCAKIYQLLGONEVDVRQKOVILLSDPSFYRLTSEOKAKALKQENPFHPDAFDNEL 169	: :
OY	97 MHRTLKINIVEKTVLEVPDYDVTSRPLETTVVYPADVLEEGILVFYSOEIRIDMFHTRL 156	: :
DB	170 ILTKLKEETBEKTYQLIPLYDVDSRKKEEITVTPADVYLEGLIAFISSQEVRDLFOVKL 229	: :
OY	157 FVDDTSDVRLSRVLRDV--RRGRDLEQILLTOYTFVFKPAFEFCPLTKKYADVII PRGVD 215	: :
DB	230 FVDTADPTRLSRVLRDISERGRDLEQLSQYITFPVKPAFEFCPLTKKYADVII PRGAD 289	: :
OY	216 NMVALNIVQHIOIILNDGIKMRGGSGNGSYKRTSPSPBDHGMILTSGRSHLESSSR 275	: :
DB	290 NLVALNIVQHIOIDLNG-----GPSKRQT--NCCLNGYTFSRRRQAESSSSR 335	: :
OY	276 PH 277	
DB	336 PH 337	
RESULT 12		
AAB56582		
ID	AAB56582 standard; Protein; 337 AA.	
XX	AAB56582;	
XX		
DT	13-MAR-2001 (first entry)	
DE		
XX	Human prostate cancer antigen protein sequence SEQ ID NO:1160.	
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;	
KM	neuroprotective; cytoskeletal; cardioactive; immunomodulatory; muscular;	
KW	vulnerable; gastrointestinal; nephrotoxic; antiinfective; gynaecological;	
KM	antibacterial; gene therapy; neural; immune; reproductive; renal;	
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;	
KX	wound; infectious disease.	
OS	Homo sapiens.	
XX		
PN	WO200055174-A1.	
XX		
DP	21-SEP-2000.	


```
XX 08-MAR-2000; 2000WO-US05988.
PF 12-MAR-1999; 99US-0124270.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM,
XX WPI; 2000-587513/55.
DR N-PSDB; AAF15785.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX Claim 11; Page 1566-1567; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardiocactive, immunomodulatory, muscular, vulnery, gaectriintestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 337 AA;
XX
Query Match 61.8%; Score 896; DB 21; Length 337;
Best Local Similarity 62.3%; Pred. No. 2,3e-78;
Matches 188; Conservative 28; Mismatches 40; Indels 46; Gaps 6;
QY 2 ASAGE-----DCESPAPADR--PHQRPFLIGVSGTASGK 36
DB 56 AGAGGERASVRTSGRRGANHGRGADPAEPAPQRRRLPYRR-----HGTASGK 109
QY 37 STYCEKIMELLIGQNEVQRQKRVILSODRFYKVLTAEQKALKQGYNDHPDPAFNDL 96
DB 110 SSVCAKIVOLLIGQNEVDYRQKVILSQDSFYRLTSEOKAKALKQFNFPHPDPAFNDX 169
QY 97 MHRTLKIVEGKTEVPTVPTVTHSRLEPTTVVPAVVLEFGLIVFSOEIRMFHLRL 156
DB 170 ILTKLKITRQKIVQIVYDFVFSRKEETVTPAVVLEFGLIARYSQEVRDLFQMKL 229
QY 157 FVPTDSVRLSRVLRDV-RRGRDLQILTOYTTFFVAPAEFCLPTKTYADVILPRGVD 215
DB 230 FVPTDADTRLSRVLDRDISERGRDLQILSQYIFVAPAEFCLPTKTYADVILPRGAD 289
QY 216 NMVAINLIYQHDIILNGDICKMHRGSGNSRYKRTSEPDHMGMLTSGKSHLESSR 275
DB 290 NLVAINLIYQHDIILNG-----GPSKRQT---NGCLNGYTPSRKRQASSESSR 335
QY 276 PH 277
DB 336 PH 337
XX
RESULT 13
ID AAM40480 standard; Protein; 335 AA.
XX
AC AAM40480;
XX
DT 22-OCT-2001 (first entry)
XX
```

```
DE Human polypeptide SEQ ID NO 5411.
XX
XX Human, nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59636.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 2; SEQ ID NO 5411; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: the sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 335 AA;
XX
Query Match 61.3%; Score 889; DB 22; Length 335;
Best Local Similarity 63.8%; Pred. No. 1,1e-77;
Matches 190; Conservative 26; Mismatches 46; Indels 36; Gaps 7;
QY 2 ASAGE-----DCESPAPADRPH-QRPFLIGVSGTASGKSTVCEK 42
DB 52 AGAGGERASVRTSGRRGRTMAGDSEQTLQNHQPNQGEPLFIVGSGTASGKSSVCAK 111
QY 43 IMELLIGQNEVQRQKRVILSODRFYKVLTAEQKALKQGYNDHPDPAFNDLMHRTLK 102
DB 112 IVOLLIGQNEVDYRQKVILSQDSFYRLTSEOKAKALKQGFNDHPDAFNDLILTKTK 171
QY 103 NIVEGKTEVPTVPTVPTVTHSRLEPTTVVPAVVLEFGLIVFSOE-IRDMFHLRLFVPTD 161
```

Db 172 ETEGKTQVQIPVYDFVSHSRKETVTVYPADVLFEGILAFYSQERIRDLFQMKLFVYD 231
QY 162 SDVLSRRVLDV-RRGDLEQILLTOYT-TVKPAFEFCFPTKKYADVILPRGVDMVA 219
Db 232 ADTLSRRVLDVSRGRDLEQILLSSSTLRFKPAFEFCFPPKKYADVILPRGADNRVP 291
QY 220 INLIVQHIQDILNDICKMRGSGNSGRSYKRTFSEPDGHPGMLTSGKSHLESSSRPH 277
Db 292 INLIVQHIQDILNG-----GPGNRQT---NGCLNGYTRSRKRGASSESSSRPH 335

RESULT 14
ABG27216
ID ABB62307 standard; Protein; 260 AA.
XX
AC ABB62307;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 13713.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR N-PSDB; ABL06410.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX interactions -
XX
PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB161840-AB16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 260 AA;
Query Match 50.9%; Score 738; DB 22; Length 260;
Best Local Similarity 67.8%; Pred. No. 3.5e-63;
Matches 143; Conservative 32; Mismatches 34; Indels 2; Gaps 2;

QY 23 PFLIGVSGTASGKSTVCEKIMELLGONEVEORQKVVILSODRFYVLTAQKAKALG 82
Db 27 PFLIGVAGTASGKSTVCKKIMEQLGQEMDHTROVVISIQDSFYELTPAEKAKOKG 86
QY 83 QYNFDHDAFNDMLMARTLKNIVGKTVVEFTYFVTHS-RLPFTTVVYPADVLFEGIL 141
Db 87 LFNFDHDAFNEELMSTLQNLKGVKVEIPSYDRTNSLDFENVLVYPADVLFEGIL 146

QY 142 VFYSOELRDMFHLRLFYDSDVLSRRVLRDV-RRGDLEQILLTOYTFVKPAFEFC 200
Db 147 VFYFPKRLRFHAKLFYDSDSLRLARVPRDINEGRDLDAVLTOYMTFVKPAFEFC 206
QY 201 PTKKPADVILPRGVDMVAINLIVQHIQDIL 231
Db 207 PTKKPADVILPRGADNTVAIDLIVHHIGEIL 237

RESULT 15
ABG27216
ID ABB27216 standard; Protein; 326 AA.
XX
AC ABB27216;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27207.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
OS WO200175067-A2.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX
XX WPI: 2001-639362/73.
XX
DR N-PSDB; AAS91403.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
PS Claim 20; SEQ ID NO 57575; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABB00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 326 AA;
Query Match 40.6%; Score 589; DB 22; Length 326;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 10:07:48 ; Search time 2539 Seconds

(without alignments)
2651.571 Million cell updates/sec

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDESPAPRADRP...HPGMLTSGKSHLESSSRPH 277

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
	Xgapop 60.0 , Ygapext 60.0
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 15

Total number of hits satisfying chosen parameters: 509

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

```
-MODE=frame+ .p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/runat_21112003_184459_5876/app_query.fasta.1.455
-DB=EST -QPM=faetap -SUFFIX=p2n_ol1.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco
-NORM=ext -HEAPSIZE=500 -MNTLEN=0 -MAXLEN=2000000000
-USER=US09896522.ccn_1.1.2810 @runat_21112003_184459_5876 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7
```

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vit.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_png.*
27: em_gss_vtl.*
28: gb_gss1.*
```

Result No.	Score	Query Match	Length	DB ID	Description
1	236	85.2	1044	12	BM917506
2	231	83.4	1201	12	BX394295
3	230	83.0	1118	12	BM467984
4	228	82.3	898	14	CD514811
5	225	81.2	725	12	B1256928
6	216	78.0	916	12	BG826894
7	214	77.3	1036	13	BQ072501
8	213	76.9	730	12	BG770518
9	213	76.9	869	10	BG491384
10	213	76.9	956	13	BQ719741
11	208	75.1	989	10	BG390519
12	205	74.0	827	12	B1261258
13	203	73.3	860	10	BG491358
14	200	72.2	1039	12	BM803129
15	197	71.1	848	12	B1258532
16	196	70.8	810	12	B1753007
17	194	70.0	1055	12	BM545603
18	193	69.7	661	12	B1830878
19	193	69.7	910	13	BQ437115
20	192	69.3	922	10	BQ935919
21	191	69.0	753	9	AU122008
22	180	68.6	963	14	CD359372
23	185	66.8	959	12	B1770572
24	185	66.8	1095	12	BM802939
25	182	65.7	778	14	CD101693
26	177	63.9	554	12	BM822058
27	174	62.8	584	10	AW961647
28	169	61.0	799	10	BG723578
29	169	61.0	1114	13	BX343101
30	168	60.6	998	13	BX400889
31	163	58.8	514	12	BM827668
32	163	58.8	515	12	BM835278
33	160	57.8	585	12	CB156232
34	160	57.8	923	10	BB891861
35	157	56.7	945	13	BQ677293
36	156	56.3	883	13	BQ838101
37	156	56.3	948	13	BQ680950
38	156	56.3	973	13	BQ683188
39	153	55.2	797	12	B1116432
40	153	55.2	898	13	BQ680187
41	152	54.9	1097	10	BF347745
42	150	54.2	819	12	B1910015
43	146	52.7	494	9	AU129365
44	145	52.3	819	12	B1770639
45	141	50.9	425	9	AI640490

ALIGNMENTS

RESULT 1
LOCUS BM917506 1044 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6606568 NIH_MGC_106 Homo sapiens CDNA clone IMAGE:5484033
ACCESSION BM917506
VERSION 5', mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1044)


```

Db      88 ATGCGCTTCGCGGAGGCGAAGACTGCGAGGCCCGCGCGGAG-GCCGACCGGCGCAC 146
Qy      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db      147 CACCGCGCCCTCTCTGATAGGGGTGAGCGCGGCGCATCTCCAGCGGGAAGTGCACCGTGTGT 206
Qy      41 GluLysIleMetGluLeuLeuGlyGlnAspGluValGluGlnArgGlyNargLysValVal 60
Db      207 GAGAGATCTAGAGTTCCTGCTGGACAGACAGAGGTGAAACAGCGCGACGCGAAGAGTGTCT 266
Qy      61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db      267 ATCTGAGCCAGAGCAGGCTCTACAGAGTCTTGAACGCGCAGAGCAGAGCGCAAGCCTTG 326
Qy      81 LysGlyGlnTyrlsAspPheAspHisProAspAlaPheAspLysPheMetHisArgThr 100
Db      327 AAGGACAGTACATTTTGAACATCCAGATGCCCTTGTATGATGATTTGATGACAGGACT 386
Qy      101 LeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrlsAspPheValThrHis 120
Db      387 CTGAGAGACATCTGAGAGGCGAAGACGTGAGAGTGCAGACTATGATTTTGTGACACAC 446
Qy      121 SerArgLeuProGluThrThrValValTyrlsProAlaAspValValLeuPheGluGlyIle 140
Db      447 TCAAGGTACACAGACACAGGTGTCTACCTGCGGACGTTGCTGTTGAGGGGATC 506
Qy      141 LeuValPheTyrlsSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      507 TTGGTGTCTACAGCGCGAGATCCGGGACATGTTCCACCTGCGCTCTGTGAGCACCC 566
Qy      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      567 GACTCGACATCGAGCTGTCTCGAAGAGTCTCCGGACGTGCGCGAGGAGGAGACTG 626
Qy      181 GluGlnIleLeuThrGlnTyrlsThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db      627 GACGAGTTCTGACGCGATCACACACCTTGTGTAAGCGCGCTTCGAGAGGTTCTGCTG 666
Qy      201 ProThrLysLysTyrlsAlaAspValIleIleProArgGlyValLysPheMetValAlaIle 220
Db      687 CCCACAAAGAAATGCGATGCGATCATCCACGAGAGTGGACATATATGTTGCCATC 746
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
Db      747 AACCTGATCGTGACGACATCCAGGACATCTGAAATGTGACATCTGCAATGACACCGA 806
Qy      241 GlyGlySerAsnGlyArg 246
Db      807 GGAGGCTCCATGGCGG 824

RESULT 3
LOCUS   BM467984 1118 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT 6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5552885
5', mRNA sequence.
ACCESSION BM467984
VERSION   BM467984.1 GI:18517026
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1118)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation

```

```

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM12216 row: n column: 22
High quality sequence stop: 689.
Location/Qualifiers
1..1118
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532885"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Nott1;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 234 a 311 c 361 g 209 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 6,47e-220 Length: 1118
Score: 230.00 Matches: 259
Percent Similarity: 99.23% Conservative: 0
Best Local Similarity: 99.23% Mismatches: 1
Query Match: 83.03% Indels: 2
DB: 12 Gaps: 0
US-09-896-522-2 (1-277) x BM467984 (1-1118)
Qy      1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluValAspArgProHis 20
Db      47 ATGCGCTTCGCGGAGGCGAAGACTGCGAGGCCCGCGCGGACCGCTCCGCGAC 106
Qy      21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db      107 CACCGCGCCCTCTCTGATAGGGGTGAGCGCGGCGCATGCGAGGAGTGCACCGTGTGT 166
Qy      41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db      167 GAGAGATCATGAGTGTCTGCGGACAGAACGAGGTGGAACAGCGCGACGCGAAGGTGTCT 226
Qy      61 IleLeuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db      227 ATCTGAGCCAGAGCAGGTTCTTACAGAGTCTGACGCGCAGACAGAGGCGCAAGGCTTG 286
Qy      81 LysGlyGlnTyrlsAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      287 AAGGACAGTACATTTTGAACATCCAGATGCTGTTGATGATGATTTGATGACAGGACT 346
Qy      101 LeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrlsAspPheValThrHis 120
Db      347 CTGAGAGACATCGTGTGAGGCGCAAAACGCGGAGGTGCCAGCTATGATTTTGTGACACAC 406
Qy      121 SerArgLeuProGluThrThrValValTyrlsProAlaAspValValLeuPheGluGlyIle 140
Db      407 TCAAGGTACACAGACACAGGTGTCTACCTGCGGAGTGTGTTGTAAGGCGACTC 466
Qy      141 LeuValPheTyrlsSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      467 TTGGTGTCTTACAGCAGAGAGATCCGGGACATGTTCCACTCGGCTCTTCTGTGACACAC 526
Qy      161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      527 GACTCGACGTCAGGCTGTCTGAAAGAGTTCCTCCGCGAGCTGCGCGGAGGAGGAGCTTG 586
Qy      181 GluGlnIleLeuThrGlnTyrlsThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db      587 GAGCAGATTTCTAGCAGAGTACACACCTTGTGAAAGCGGCTTGTGAGGAGTTCTGCTGTG 646
Qy      201 ProThrLysLysTyrlsAlaAspValIleIleProArgGlyValAspAspMetValAlaIle 220
Db      647 CCGACAAAGAAATGAGCGAGTGTGATCATCCACGAGAGGTGGAATATGTTGTTGCCATC 706

```

Qy 221 AsnLeuIleValGlnHisIleGlnAspIle-LeuAsnGlyAspIleCysIleTyrPheSar 240
 Db 707 AACCTGATCGTGCAGCAGCATCCAGACATTTCTGAT-GGTGACATCTGCAGAAATGGCACC 765
 Qy 240 GGLGlySerAsnGlyThrSerSerTyrLysArgThrPheSerGlyProGlyAspHisProGly 260
 Db 766 AGAGGGGTCCATGGCGGAGCTCAACCGGACCTTTTCTGAGCCAGGGAGCCACCTCG 825
 Qy 260 Y 260
 Db 826 N 826
 RESULT 4
 CD514811 898 bp mRNA linear EST 06-JUN-2003
 LOCUS AGNCCOURT 14376196 NIH_MGC_181 Homo sapiens cDNA clone
 DEFINITION IMAGE:30396201 5', mRNA sequence.
 CD514811
 ACCESSION CD514811.1 GI:31446529
 VERSION EST.
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE
 REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM477 row: i column: 10
 High quality sequence stop: 662.
 FEATURES
 source location/Qualifiers
 1..898
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30396201"
 /tissue_type="White Matter"
 /dev_stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_id="NIH_MGC_181"
 /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
 (destroyed); Library is oligo-dr primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.42 kb. Library was constructed by
 (Invitrogen). Note: this is a NIH_MGC Library."
 BASE COUNT 194 a 231 c 310 g 158 t 5 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.18e-218 Length: 898
 Score: 228.00 Matches: 228
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 82.31% Indels: 0
 DB: 14 Gaps: 0
 US-09-896-522-2 (1-277) x CD514811 (1-898)
 Qy 1 MetAlaSerAlaGlyGlyGlyAspCysGlySerProAlaProGlyAlaAspArgProHis 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 72 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGGAGGCGCAGCCGCCGAC 131
 Qy 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 132 CAGCGGCCCTCTCTGATGAGGGGTGACCGCGGCGACCGCAGCGGAAGTCGACCCGTGT 191
 Qy 41 GlnLysIleMetGlnLeuLeuGlyGlnGlnGlnValGlnGlnValArgGlnArgLysValVal 60
 Db 192 GAGAGATCATGAGATTGCTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251
 Qy 11LeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
 Db 252 ATCTTGAGCAGGAGACAGGTTCTTCAAGGTCCTGACCGCAGAGAGAGAGAGAGAGAG 311
 Qy 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 312 AAAGGACAGTACATTTTGCATCCAGATGCTTTGATATGATTTGATGACAGAACT 371
 Qy 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
 Db 372 CTGAGAGACATCGTGAAGGAGCAAAACGTTGAGAGTCCAGCTATGATTTTGTGACACAC 431
 Qy 121 SerArgLeuProGlyThrThrValValTyrProAlaAspValLeuPheGlnGlyIle 140
 Db 432 TCAAGGTACCAAGACACAGCGGTGTCTACCTGCGGAGCGTGTCTGTTGAGGGCATC 491
 Qy 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 492 TTGGTGTTCACAGCGCAGAGATCCGAGCATGTTCCACTGCGCTCTTCGGGACAC 551
 Qy 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 552 GACTCCGACGTGAGCTGTCTGCAAGAGTTCTCCGGAGAGTCCGCGAGGAGGAGCTTG 611
 Qy 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnPheCysLeu 200
 Db 612 GACGAGTTCTGACGAGTACACACCTTCGTGAACCGGCTTCGAGGAGTTCTGCTCC 671
 Qy 201 ProThrLysLysTyrLysAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 672 CCGACAAAGAAATATCCGATGATCATCCACGAGAGTGAACAATATGTTGTCATC 731
 Qy 221 AsnLeuIleValGlnHisIleGln 228
 Db 732 AACCTGATCGTGCAGCAGCATCCAG 755
 RESULT 5
 BI256928 725 bp mRNA linear EST 17-JUL-2001
 LOCUS 602975534F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114816 5',
 DEFINITION mRNA sequence.
 BI256928
 ACCESSION BI256928.1 GI:14811813
 VERSION EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LHAM1279 row: c column: 09
 High quality sequence stop: 707.


```

FEATURES
    source
        location/Qualifiers
            1..725
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5114816"
                /tissue_type="cervical carcinoma cell line"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_12"
                /note="Organ: Cervix; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
BASE COUNT      175 a      181 c      224 g      145 t
ORIGIN
Alignment Scores:
Pred. No.:      4,19e-215      Length:      725
Score:          225.00      Matches:      225
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    81.23%      Indels:      0
                Gaps:      0
US-09-896-522-2 (1-277) x B1256928 (1-725)
QY      36  LysSerThrValCysGluValIleMetGluLeuGluGlnAangluValGluGlnArg 55
DB      3  AAGTCGACCGTGTCGAGAGATCATGAGTTGCTGGACAGAACGAGGTGGAAACACGGG 62
QY      56  GluArgValValIleleuSerGlnAspArgPheTyrLysValIleuThrAlaGluGln 75
DB      63  CAGCGAAGGTGTCATCTGAGCCAGAGACGGTTCTACAAAGTCTTGACGCGACAGACGAG 122
QY      76  LysAlaLysAlaLeuLysGluGlnTyrAsnPheAspIleProAspAlaPheAspAsnAsp 95
DB      123  AAGGCCAAGGCTTGAAAGGACAGTACATTTTACCATTCAGATGCTTTGATTAAGAT 182
QY      96  LeuMetIleSarGThrLeuLysAsnIleValGluGluValThrValGluValProThrTyr 115
DB      183  TTGATGACACGAGACTCGAAGAACATGCTGAGGCGCAAAACGGTGGAGTGCACCACTAT 242
QY      116  AspPheValThrIleSerArgLeuProGluThrThrValValTyrProAlaAspValVal 135
DB      243  GATTTTGACACACTAGAGGTACCAAGACCAACGGTGTCTACCTCGGAGACGTGAT 302
QY      136  LeuPheGluGluIleLeuValPheTyrSerGlnGluIleArgAspMetPheIleLeuArg 155
DB      303  CTGTTTGAGGGCATCTTGATGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTGCGC 362
QY      156  LeuPheValAspThrAspSerAspValArgLeuSerArgArgValIleuArgAspValArg 175
DB      363  CTCCTTCGTGACACCGACTCGAGAGTCAAGCTGCTCGAAGAGTTCCTCCGGGACGTCCG 422
QY      176  ArgGluArgAspLeuGluGlnIleLeuThrGlnThrThrPheValLysProAlaPhe 195
DB      423  CGAGGAGGAGGACCTCGAGACGATTTCTACGCGATTCACCACTTCGTAAGACCCGCGCTTC 482
QY      196  GluGluPheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGluValAsp 215
DB      483  GAGGAGTTCGCTGCGCCGACAAAGAGATATCCGATGTGATCATCCACAGAGAGTGGAC 542
QY      216  AsnMetValAlaIleAsnLeuIleValGlnIleIleGlnAspIleLeuAsnGluAspIle 235
DB      543  AATATGTTGTCATCAACCTGATCGTCAAGACATCAAGAGCATTCGAATGTGATCATC 602
QY      236  CysLysTrpHisAspGlyGlySerAsnGlyLysArgSerTyrLysArgThrPheSerGluPro 255
DB      603  TGCAAATGGACCCAGAGAGGATCCAAATGGGCGAGCTACAAAGCGAACCTTTCTGAGCCA 662
QY      256  GlyAspHisProGly 260
DB      663  GGGGACCACTGGG 677

```

```

RESULT 6
BG826894
LOCUS      916 bp      mRNA      linear      EST 22-MAY-2001
DEFINITION 602750978F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
mRNA sequence.
ACCESSION  BG826894
VERSION    BG826894.1  GI:14174481
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 916)
            NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished
JOURNAL   Contact: Robert Strausberg, Ph.D.
COMMENT   Email: csabpe-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LNCMI802 row: k column: 20
            High quality sequence stop: 843.
FEATURES
    source
        location/Qualifiers
            1..916
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4903819"
                /tissue_type="rhabdomyosarcoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_17"
                /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      203 a      239 c      309 g      165 t
ORIGIN
Alignment Scores:
Pred. No.:      5.73e-206      Length:      916
Score:          216.00      Matches:      216
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    77.98%      Indels:      0
                Gaps:      0
US-09-896-522-2 (1-277) x BG826894 (1-916)
QY      1  MetAlaSerAlaGlyGluGluSerGluSerProAlaProGluAlaAspArgProHis 20
DB      80  ATGGCTTCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 139
QY      21  GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB      140  CAGCGGCGCTTCCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 199
QY      41  GluValIleMetGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 60
DB      200  GAGAAGATCTGAGAGTTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 259
QY      61  IleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGluGlnLysAlaLysAlaLeu 80
DB      260  ATCTTAGGCGAGAGAGGTTCTACAAAGTCTTACGCGCAGAGAGAGAGAGAGAGAGAGAG 319

```

```

Qy      81  LysGIyGlnTYrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      320 AAAGACAGCTACATTTTGACCATCCAGATGCTTTGTAATGATTTGATCAGACGACT 379
Qy      101  LeuLYsAsnIleValGlnGlyLYrThrValGluValProThrTYrAspPheValThrHis 120
Db      380 CTGAAGAATCATCTGTGAGGGGCAAAACGCTGAGAGGTCCGACCTATGATTTGTGACACAC 439
Qy      121  SerArgLeuProGlnThrThrValValTYrProAlaAspValValLeuPheGlnGlyIle 140
Db      440 TCAAAGTTACAGACAGACACGCTGCTACCCCTGGACGGGTCTGTGTTGAGGGCACTC 499
Qy      141  LeuValPheTYrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      500 TTGGTGTTCTACAGCCAGACGATCCGGAGACATGTTCCACCTGCGCTCTTGTGTGACACC 559
Qy      161  AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      560 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCGCCGAGGAGGAGACTG 619
Qy      181  GlnGlnIleLeuThrGlnTYrThrThrPheValLYrProAlaPheGlnGlnPheCysLeu 200
Db      620 GAGCAGATTCTGACGCGACATACACACCTTCGTGAAGCCGCGCTTGAAGAGATTCTGCTG 679
Qy      201  ProThrLYrLYrTYrAlaAspValIleIleProArgGlyValAspAsn 216
Db      680 CCGAACAAAGATATGCCGATGTATCATCCACGAGAGGTGACAAAT 727

RESULT 7
LOCUS   BQ072501 1036 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT 6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694
5', mRNA sequence.
ACCESSION BQ072501
VERSION   BQ072501.1 GI:19901547
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1036)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgs@bbs-remail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNLN at:
           http://image.llnl.gov
           Plate: LLM12810 row: 1 column: 15
           High quality sequence stop: 644.
           Location/Qualifiers
             1..1036
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5761694"
               /lab_host="DH10B"
               /clone_1lb="NIH MGC 122"
               /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
               Site_1: NoCI; Site_2: EcoRV (destroyed); RNA source
               anonymous pool of 24 week female lung, 16 week female
               spleen, and 20-22 week male spleens. Library is oligo-dT
               primed and directionally cloned (EcoRV site is destroyed
               upon cloning). Average insert size 1.4 kb. Insert size
               range 1-3 kb. Library is normalized and enriched for
               full-length clones and was constructed by C. Gruber
               (Invitrogen). Research Genetics tracking code 026. Note:

```

```

BASE COUNT      221 a      285 c      346 g      184 t
ORIGIN
Alignment Scores:
Pred. No.:      6.63e-204      Length:      1036
Score:          214.00      Matches:      227
Percent Similarity: 99.56%      Conservative: 0
Best local Similarity: 99.56%      Mismatches: 0
Query Match:    77.26%      Indels:      1
DB:             13      Gaps:      0

US-09-896-522-2 (1-277) x BQ072501 (1-1036)
Qy      1  MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGlnAlaAspArgProHis 20
Db      64 ATGGCTTTGGCGGAGGCGAAGACTGCGAAGCCCCCGCGGAGGCCGACCTCCGAC 123
Qy      21  GlnArgProPheLeuIleGlyValSerGlyTYrAlaSerGlyLYrSerThrValCys 40
Db      124 CAGGGGCCCTTCTGATAGGGGTGAGCGGCGCACTGCCAGCGGGAAGTGAACGTTGT 183
Qy      41  GlnLYrIleMetGlnLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgLYrVal 60
Db      184 GAGAGATCATGTGAGATTGCTGGGACAGAAACGAGTGGAAACAGCGCAGCGGAAGTGTG 243
Qy      61  IleLeuSerGlnAspArgPheTYrLYrValLeuThrAlaGlnGlnLYrAlaLYrAlaLeu 80
Db      244 ATCTTGAGCCAGACACAGGTTCTTCAGAGTCTTACCGGACAGACAGAACGCCAGCTTG 303
Qy      81  LysGIyGlnTYrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      304 AAAGACAGTACATTTTGACATCCACAGATGCTTGTATATATGATTTGATGACAGAGACT 363
Qy      101  LeuLYsAsnIleValGlnGlyLYrThrValGluValProThrTYrAspPheValThrHis 120
Db      364 CTGAAGAATCATCTGTGAGGGGCAAAACGTGGAGGTCCGACCTATGATTTGTGACACAC 423
Qy      121  SerArgLeuProGlnThrThrValValTYrProAlaAspValValLeuPheGlnGlyIle 140
Db      424 TCAAAGTTACAGACAGACACGAGTGTCTTACCTTCGCGAGCTGTGTTTGAGGGCATC 483
Qy      141  LeuValPheTYrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      484 TTGGTGTTCTACAGCCAGACGAGATCCGGACATGTCACCTGCGCTTCTGTGACAC 543
Qy      161  AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db      544 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGTGCGCGAGGAGGACCTG 603
Qy      181  GlnGlnIleLeuThrGlnTYrThrThrPheValLYrProAlaPheGlnGlnPheCysLeu 200
Db      604 GAGCAGATTCTGACGACAGTACACCACTTCGTGAACCGGCTTCCAGAGAGTTCTCCTG 663
Qy      201  -ProThrLYrLYrTYrAlaAspValIleIleProArgGlyValAspAsnMetValAlaI 220
Db      664 CCGACAAAGATATGCCGATGTGATCATCCACAGAGGTGACAAATATGTTGCCAT 723
Qy      220  eAenLeuIleValGlnHisIle 227
Db      724 CAACCTGATGTCAGACACATC 745

RESULT 8
LOCUS   BG770518 730 bp mRNA linear EST 15-MAY-2001
DEFINITION BG770518 602734236F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4859577 5',
mRNA sequence.
ACCESSION BG770518
VERSION   BG770518.1 GI:14081171
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
1 (bases 1 to 730)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LICM1715 row: h column: 10
High quality sequence stop: 720.
Location/Qualifiers

FEATURES
source

1..730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4859577"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 49"
/note="Organ: skin; Vector: pOTB7, site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 156 a 190 c 258 g 126 t
ORIGIN

Alignment Scores:

Pred. No.: 4,57e-203 Length: 730
Score: 213.00 Matches: 213
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.90% Indels: 0
Gaps: 0

US-09-896-522-2 (1-277) x BG770518 (1-730)

QY 1 MetAaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 82 ATGCTTCGGGGGAGGAGGAGAGCTGCGAGAGCCCGCGCGAGCGACCGCCGAC 141
QY 21 GlnArgProPheLeuIleGlyValSerGlyValThrAlaSerGlyLysSerThrValCys 40
DB 142 CAGCGGCCCTTCCTCGATAGGGGTGAGCGCGCGCATCCAGCGGGAAGTCACCGTGT 201
QY 41 GluValIleMetGluLeuGluValGlnAsnGluValGluGlnArgGlnArgGlyValVal 60
DB 202 GAGAGAGTCATGAGGTGCGGACAGAAACAGAGTGAACCGCGCAGCGAAGGTGCTC 261
QY 61 IleLeuSerGlnAspArgPheTyrGlyValLeuThrAlaGluGlnValAlaValAlaLeu 80
DB 262 ATCTGAGCCAGAGCAGAGTTCACAGTCTCTGACGCGAGAGCAGAAAGCCAGCTTG 321
QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 322 AAAGAGACAGTAACTTTGACCATCCAGATGCTTTATATGATTTGATGACAGACT 381
QY 101 LeuValAsnIleValGluGlyValThrValGluValProThrTyrAspPheValThrHis 120
DB 382 CTGAAGAACATCGAGAGGGGCAAAACGATGAGGTGCGACCTATGATTTTGACACAC 441
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140

DB 442 TCAAGGTACAGAGACACAGGTGTCTACCTCGCGAGCGTGGTCTGTTCAGGCAATC 501
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 502 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCACCTGCGCTTCTGTGGACACC 561
QY 161 AspSerAspValArgLeuSerArgValLeuAlaAspValArgArgGlyArgAspLeu 180
DB 562 GACTCCGACACTCAGCGTGTCTCGAAGATTCTCCGGGACCTGCGCGAGGAGGACCTG 621
QY 181 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 622 GACAGATTCTGACGACATACACACCTTCGTGAACCGGCTTCAGAGAGTCTGCTG 681
QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGly 213
DB 682 CCGACAAAGATGATGCCGATGTGATCATCCACGAGGA 720

RESULT 9
LOCUS BG491384 869 bp mRNA linear EST 27-MAR-2001
DEFINITION 602535670F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:484936 5', mRNA sequence.
ACCESSION BG491384
VERSION BG491384.1 GI:13452896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LICM1493 row: k column: 17
High quality sequence stop: 863.
Location/Qualifiers

FEATURES
source

1..869
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:484936"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 41"
/note="Organ: skin; Vector: pOTB7, site 1: XhoI; site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 197 a 223 c 291 g 156 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 5.52e-203 Length: 869
Score: 213.00 Matches: 239
Percent Similarity: 99.17% Conservative: 0
Best Local Similarity: 99.17% Mismatches: 0
Query Match: 76.90% Indels: 2
Gaps: 0

US-09-896-522-2 (1-277) x BG491384 (1-869)

QY	1	MetAlaSerAlaGlyGlyGlnuAspGlyGluSerProAlaProGluAlaAspArgProHis	20
Db	52	ATGGCTTGGCGGGGCGGAGACATCGCGAGCCCCGGCGGAGAGCCGACGCTCGCAC	111
QY	21	GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys	40
Db	112	CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCCGACTGCACGCGGAGATCGACCTGTGT	171
QY	41	GluLysIleMetGluLeuLeuGlyGlnuAsnGlyValGluGlnArgGlnArgLysValVal	60
Db	172	GAGAAAGATCATGAGCTGCTGSGAAGAAACGAGGTGGAAACGCGGACGAGAAAGGTGTCT	231
QY	61	IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysValAlaLeu	80
Db	232	ATCTCGAACCCAGGACAGGTTCTACAAAGTCTCGACGGCGAGCAAAAGCCAAAGCCCTTG	291
QY	81	LysGlyGlnTyrAsnPheAspHisPProAspAlaPheAspAsnAspLeuMetHisArgThr	100
Db	292	AAAGGACAGTCAATTTTGACCATCCAGATGCTTTGATTAATGATTTATGATGACAGGACT	351
QY	101	LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheAlaHis	120
Db	352	CTGAAGAACATGCTGTGAGGGCAAAACGGTGGAGGTGCCGACCTATGATTTTGTGACAC	411
QY	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe	140
Db	412	TCAAGGTTACAGAAACACACGCGTGTCTACCTCGCGAGCGTGTCTGTTTGAGGGCATTC	471
QY	141	LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
Db	472	TTGTGTTCTACAGCCAGGAGATCCGGGACATGTTCACCTCGCGCTCTTGTGTGACACC	531
QY	161	AspSerAspValArgLeuSerArgThrValLeuArgAspValArgArgGlyArgAspLeu	180
Db	532	GACTCCGACGTCACGGCTGCTCTGGAAGAGTTCCTCGGAGCGTGCCTCGGAGGAGGACCTG	591
QY	181	GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheValLeu	200
Db	592	GAGCAGGATTCGACCCAGACACCACTTCGGAAGCGCGCTTGGAAGAGTTCGCTCG	651
QY	201	ProThrLys-LysTyrAlaAspValIleIlePProArgGlyValAspAsnMet-ValAlaI	220
Db	652	CCGACMAAGAAAGTATGCCGATGTGATCATCCACAGAGAGTGGACAAATATGGGTGCCA	711
QY	220	IleAsnIleIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPHis	239
Db	712	TCAACTGATCGTGGACGACATCCAGGACATTTCTGAATGGTGACATTCGCAATGGCAC	770

REFERENCE
1 (bases 1 to 956)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE	JOURNAL	COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished	Contact: Robert Straubeberg, Ph.D. Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski		
CDNA Library Preparation: Life Technologies, Inc.		
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
DNA sequencing by: Agencourt Bioscience Corporation		
Clone distribution: MGC clone distribution information can be		

Found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1AM13574 row: b column: 18
 High quality sequence stop: 552.

FEATURES	Location/Qualifiers
source	1. .956

BASE COUNT	202 a	272 c	296 g	186 f
ORIGIN				

Alignment Scores:

Pred. No.:	6.12e-203	Length:	956
------------	-----------	---------	-----

Score:	213.00	Matches:	21
Percent Correct:	100.00%	Consecutive:	0

Percent similarity:	100.00%	Conservative:	0
Best local similarity:	100.00%	Mismatches:	0

```
Query Match: 76.90% Indels: 0
```

DB:	13	Gaps:	0
-----	----	-------	---

PAGE NO 006 E33 3 (1-277) -- PAGE08741 (1-056)

	US-09-896-522-2 (1-277) x BQ179741 (1-956)
Qy	6 GYGLuAsPCySGlUSeProAlaProGluAlaAspArgProHisGlnArgProPheLeu 25
Db	22 GGGGAAGACTGCAGAGCCCGCGCGGAGGCCGACCTGCCACACGCGCCCTTCCTG 81
Qy	26 IIEGIVAlSeScIyGLIYThraLaSeSrgIyLSeSeThraIVaSVeGluIyStIeWeGlu 45
Db	82 ATAGGGGTGAGCGCGCGCACTGCGACGGGGAAGTGACCGGTGTGGAAAGATCATTGAG 14
Qy	46 IeUleuGIVGlnAsmIuValGluGlnArgGlnArgIyValIleIeUSeScIuAsp 65
Db	142 TTGCTGGGACAGAACAGAGTGGACAGCGCGACCGGAAGTGTATCTTACACGAGAC 20
Qy	66 ArgPheTyrIyValIeUThraIaGluGlnIySalIyValIeUleuGIVGlnTyrAsn 85
Db	202 AGGTTCACAAAGGTCCTGACGGCAGAGCAGAAAGCCAAAGCCCTTGAAAGACAGTACAT 26
Qy	86 PheAspHisProAspAlaPheAspAsnAspLeuWeCHIsaGlyThrLeuIyAsnIleVal 10
Db	262 TTTCACCATTCACAGATCCCTTGATATATGATTTTGACACACAGACCTCAAGACATCTGTG 32
Qy	106 GluGluIyLysThrValGluIuValProThrTyrAspPheValIThrHisSerArgLeuProGlu 12
Db	322 GAGGGCAAAACGGTGGAGGTGGCGACCTATGATTTTGACACACTCAAGGTACCGAG 38
Qy	126 ThrThrValValTyrProAlaAspValIleUleuPheGluGluIleUleuValPheTyrSer 14
Db	382 ACCACGGTGTCTACCCCTGCGAGCTGTGTCTGTTGAGGCGATTTGTGTCTACAGC 44
Qy	146 GlnGluIleArgAspPhePheHisIeUArgLeuPheValAspThrAspSerAspValArg 16
Db	442 CAGGAATCCGGGACATGTTCACCGCGGCGCTTCGTGGAACACGACCTCGGAGCTCAG 50
Qy	166 LeuSerArgArgValIleuArgAspValArgArgGlyArgAspLeuGluGlnIleUleuThr 18
Db	502 CTGTCTCGAAGAGTTCTCCGGACGTGGCCCAAGGAGAGGACCTGGACACAGATTCGACG 56

Oy	186	GlnYrThrThrpPheValValProAlaIahGedInGubPheCysIeuProThrlYbLysYr	205
Db	562	CAGRACACACCTTCGTGTGAAGCCGCCGCTTCAGAGGTTCTCTGCCACAAAGATAT	621
Oy	206	AlaSpValIleIleIleProArgGlyValAspAsnMetVal	218
Db	622	CCCGATGTGATCATCCCGCGAGAGAGTGTGACAAATATGTT	660
RESULT 11			
LOCUS	BG390519	989 bp	mRNA linear EST 12-MAR-2001
DEFINITION	60241618371 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524308 5',		
	mRNA sequence.		
ACCESSION	BG390519		
VERSION	BG390519.1	GI:13283967	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Eumetazoa; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1 (bases 1 to 989)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D.		

```

Plate: LLAM10428  row: 5  column: 21
High quality sequence stop: 750.
Location/qualifiers
1. .989
FEATURES
source

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4524308"
/issue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH MGC 92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."

```

BASE COUNT	256 a	258 c	314 g	161 t
ORIGIN				
Alignment Scores:				
Pred. No.:	6.56e-198		Length:	989
Score:	208.00		Matches:	208
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	75.09%		Indels:	0
DB:	10		Gaps:	0

US-09-896-522-2 (1-277) x BG390519 (1-989)

7 **GI**Asp**PC**ys**GI**ser**PR**o**Al**a**PR**o**GI**u**A**l**As**p**Arg****PR**o**HI**s**GI**na**Arg****PR**o**PH**e**Le**u**I**e 26

Db 2 GAAGACTGGAGAGCCCCCGCCGGAGGCCGACCGTCCGCACCAAGCGCCCTTCTGATA 61

27 GYVAISERIGLYINRAASERIGLYSSEIRINRAVALCYSGULYSIEMETGULEU 46

[illegible]

Db	122	CTGGGACAGAAACGAGGTGGAAACAGCGGACGGACGGAAGTGGTCAATCTTGAGCCAGGACAGG	181
QY	67	PheTYrIlySValLeuThrIalaglInLysAlaLysAlaLeuIlysgIyGlnTYrAsn	86
Db	182	TTCTACAGAGGCTCCGAGCGGACGACAGAGCCAAAGGCTTGAAAGGACAGTCAATTTT	241
QY	87	AspIAspProAspAlaPheAspAsnAspLeuMetIAsrGThrLeuLysAsnIleValGlu	106
Db	242	GACCATCCAGATGGCTTTGATTAAGATTGTATGCACAGCACTCGAAGAACATGTGTGAG	301
QY	107	GlyLysThrValGluValProThrTYrAspPheValThrIIsSerArgLeuProGluThr	126
Db	302	GGCAAAACGGGGAGGGGCGGACCTTAAGATTGTTGTGCACACTCAAGGTTACAGAGACC	361
QY	127	ThrValValTYrProAlaAspValValLeuPheGluValIleLeuValPheTYrSerGln	146
Db	362	ACGGTGGTCTACCTCGCGGAGCGTGGTCTCTGTTGGAGGCACTTTGGTGTCTACAGCCAG	421
QY	147	GluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgLeu	166
Db	422	GAGATCCGGGACATGTTCCACCTGCGGCTCTTCGTGGACACCGCATCCGACGTACAGGCTG	481
QY	167	SerArgArgValLeuArgAspValAlaArgGlyAlaArgAspLeuGlnGlnIleLeuThrGln	186
Db	482	TCTGAGAGAGTTCTCCGGAGCGTGGCGGAGGGAGGGAAGCTGGAGCAGATTTCGACCCAG	541
QY	187	TYrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLysLysTYrAla	206
Db	542	TACACCAACCTTCGTGAAGCGGCGCTTCGAGGAGTTCTGCTCGGACCAAGAGATATGCC	601
QY	207	AspValIleIleLeuProArgGlyVal	214
Db	602	GATTTGATTCATCCGCGAGGAGTT	625

RESULT 12
BI261258
LOCUS
DEFINITION
602965110F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5108666 5',
mRNA sequence.
ACCESSION
BI261258
VERSION
BI261258.1 GI:14820349
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 827)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished
JOURNAL
Contact: Robert Strausberg, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Plate: L14M1263 row: C column: 03
High quality sequence start: 5
High quality sequence stop: 799.

FEATURES	Location/Qualifiers
source	1. .827

```

/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/ad_xref="taxon:9606"
/clone="IMAGE:5108666"
/tissue_type="cervical" carcinoma cell line"

```

```
/add_nosrc="DHIOB"  
/clone_11b="NTH MGC 12"
```

Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT 189 a 210 c 274 g 154 t

ORIGIN

Alignment Scores:

Pred. No.: 5.52e-195 Length: 827
Score: 205.00 Matches: 218
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 74.01% Indels: 1
DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x B1261258 (1-827)

Qy 1 MetAlaSerAlaGlyGlyGluAAspCyGluSerProAlaProGluAlaAspArgProHis 20
Db 52 ATGGCTTCGGGGGGGGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGCAC 111
Qy 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 112 CAGGGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCCAGCTGTGT 171
Qy 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
Db 172 GAGAGATCATGAGATTCTCGGACAGAAACAGAGTGGAAACGGCGCAGCGAAGGTGTC 231
Qy 61 IleLeuSerGlnAspArgPheTyrlValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 232 ATCCGAGCCAGGACAGGTTCTTCAAGAGTCTTGACGCGAGAGCGAAGGCCCTTG 291
Qy 81 LysGlyGlnTyranPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 292 AAAGGACAGTCAATTTTGCATCCAGATCCCTTGTATGATGATTCACAGAGACT 351
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThyTyranPheValThrHis 120
Db 352 CTGAAGAACATCGTGAGAGCGCAAAACGGTGGAGGCGGACCTATGATTTGTGACACAC 411
Qy 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
Db 412 TCAAGGTTACAGAGACACAGGTGTCTCACTCCGCGGACGGTGTCTGTTGAGGGATC 471
Qy 141 LeuValPheTyranPheSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 472 TTGGTGTTCACACCGCAGAGATCCGGACATGTTCCACCTCCGCTCTTGTGTGACACC 531
Qy 161 AspSerAspValArgLeuSerArgArgValLeu-ArgAspValArgArgGlyArgAspLeu 180
Db 532 GACTCCGACGTCAGGCTGTCTCGAAAGAGTCTCCCGGAGCGTGGAGGAGGAGACT 591
Qy 180 uGluGlnIleLeuThrGlnTyrrThrPheValLysProAlaPheGluGluPheCysLe 200
Db 592 GGAGCAGATTCTGACGCGAGTACACCACTTGTAAAGCGGCGCTTCAGAGGTTCTGCT 651
Qy 200 uProThyLysLysTyranAlaAspValIleIleProArgGlyValAspAsnMetVal 218
Db 652 GCCGACAAAGAGATGTCGCGATGTGATCATCCACGAGAGTGGACATATAGTT 706

RESULT 13
BG491358 860 bp mRNA linear EST 27-MAR-2001
LOCUS 60253642P1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684738 5',
DEFINITION mRNA sequence.
ACCESSION BG491358
VERSION BG491358.1 GI:13452870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS NIH-MGC <http://mhc.ncbi.nlm.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM493 row: c column: 11
High quality sequence stop: 844.
Location/Qualifiers

FEATURES
source 1..860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4684738"
/issue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into BcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC library."

BASE COUNT 195 a 223 c 288 g 154 t

ORIGIN

Alignment Scores:
Pred. No.: 5.83e-193 Length: 860
Score: 203.00 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 73.29% Indels: 1
DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x BG491358 (1-860)

Qy 1 MetAlaSerAlaGlyGlyGluAAspCyGluSerProAlaProGluAlaAspArgProHis 20
Db 52 ATGGCTTCGGGGGGAAGACTGCGAGAGCCCGCGGAGCGGACCGTCCGCAC 111
Qy 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
Db 112 CAGCGGCCCTTCTTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTCCAGCTGTGT 171
Qy 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
Db 172 GAGAGATCATGAGATTCTCGGACAGAAACAGAGTGGAAACGGCGCAGCGAAGGTGTC 231
Qy 61 IleLeuSerGlnAspArgPheTyrlValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 232 ATCCGAGCCAGGACAGGTTCTTCAAGGTCCTGACCGCAGAGAGAGCCAGGCTTG 291
Qy 81 LysGlyGlnTyranPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 292 AAAGGACAGTCAATTTTGCATCCAGATCCCTTGTATGATGATTTGATGACAGAGACT 351
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThyTyranPheValThrHis 120
Db 352 CTGAAGAACATCGTGAGAGCGCAAAACGGTGGAGGTCGACCTATGATTTTGGACACAC 411
Qy 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
Db 412 TCAAGGTTACAGAGACACAGGTGTCTTACCTTCGCGGACGTGTTCTGTTGAGGGATC 471
Qy 141 LeuValPheTyranPheSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160

Db	472	TTGGTGTTCTACAGCCAGAGAGTCCGGAGACATGTTCCACCTCGCGCTCTTGTTGGACACC	531
Qy	161	AspergerApVal1AaGLEuSerIrgArg-ValleuAArgApVal1AArgaGlgValAaPle	180
Db	532	GACTCCGACGTCACGGCTGCTCTCGAAGAAGTTCTCCGGGACGTCGCCGAGGAGGAGACT	591
Qy	180	uGluGlnIleuPThrgInrYrThThaPhaValysProAlaPheGluGluPheCysLe	200
Db	592	GGAGCAGATCTCGACGACGATACACCACTTCGTGAAGCCGGGCTTCGAGAGTTCTGCT	651
Qy	200	uProThrIysIysTYrAlaAspVal1Ile1EProArGlgVal1AspAan	216
Db	652	GCCGACAAAGAAGTATGCCGATGTCATCATCTCCACGAGAGTGGACAT	700
RESULT 14			
LOCUS	BM803129	1039 bp	mRNA linear EST 05-MAR-2002
DEFINITION	AGENCOURT 6468837 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5721513		
ACCESSION	BM803129		
VERSION	BM803129.1	GI:19119952	
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1039) NIH-MGC http://mgs.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished		
AUTHORS	Contact: Robert Strusberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Invitrogen		
TITLE	cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation		
JOURNAL	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LLM12706 row: b column: 10 High quality sequence stop: 652.		
COMMENT	Location/Qualifiers 1. 1039 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5721513" /lab_host="VDH10B" /clone_lib="NIH_MGC_125" /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool of three ovaries, from females ranging in age from 38 to 49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."		
FEATURES			
SOURCE			
BASE COUNT	240 a	266 c	309 g 220 t 4 others
ORIGIN			
Alignment Scores:			
Pred. No.:	7.3e-190	Length:	1039
Score:	200.00	Matches:	200
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	72.20%	Indels:	0
DB:	12	Gaps:	0
Qy	78	LysAlaLeuLysGlyGlnTyrAsnPhaSpHsPProAlaPheAspAanAspLeuMet	97

[illegible]

/clone_11b="NIH_MGC_12"
/note="Organ: cervix; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 196 a 219 c 271 g 162 t
ORIGIN

Alignment Scores:

Pred. No.:	5.98e-187	Length:	848
Score:	197.00	Matches:	210
Percent Similarity:	99.53%	Conservative:	0
Best Local Similarity:	99.53%	Mismatches:	0
Query Match:	71.12%	Indels:	1
DB:	12	Gaps:	0

US-09-896-522-2 (1-277) x BI258532 (1-848)

```
QY 37 SerThrValCysGluLysIleMetGluLeuGluGlnAsnGluValGluGlnArgGln 56
Db 87 TCGACCGGTGTGAGAGATCATGAGTGTCTGGGACAGACGAGTGGACACGGCGCAG 146
QY 57 ArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLys 76
Db 147 CGGAGGTGTCTCATCTGAGCAGACAGAGTTCACAGGTCTGACGCGCAGACAGAG 206
QY 77 AlaLysAlaLeuLysGluGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 96
Db 207 GCCAAGGCTTGAAGACATGACATTTTTCACATCCAGATGCCCTTTGATATGATTTG 266
QY 97 MethIAsnGlnLeuLysAsnIleValGluGlnLysThrValGluValProThrTyrAsp 116
Db 267 ATGCACAGACTCTGAAGACATCATGTGGAGGCAAAACGGTGGAGGTGCCGACCTATGAT 326
QY 117 PheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeu 136
Db 327 TTTGTGACACCTCAAGTTACAGAGCCACGGTGTCTACCTCGGACGTGTTCTG 386
QY 137 PheGluGlyIleLeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeu 156
Db 387 TTTGAGGGCATCTTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTC 446
QY 157 PheValAspThrAspSerAspValArgLeuSerArgValLeuArgAspValArgArg 176
Db 447 TTCGTGACACCGACTCGACGTCAAGGTCTCGAAGAGTTCTCCGGAGCTGCGCGA 506
QY 177 GluArgAspLeuGluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGlu 196
Db 507 GGGAGGAGCTCGAGCATTTCTGACGACATCACACCTTCGTGAAGCCGGCCTTCGAG 566
QY 197 GluPheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsn 216
Db 567 GAGTTCGTGCTCGCGACAAAGATATGCCATGTGATCATCCACAGAGAGTGGACAT 626
QY 217 MetValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsn-GlyAspIleCys 236
Db 627 ATGGTGCATCAACCTGATCGTGCACACATCCAGGACATTTGAAATGGTGCATCTG 686
QY 236 GlyTyrHisArgGlyLysSerAsnGlyArg 246
Db 687 CAATGGACCGAGAGGAGGTCATATGGGCGG 717
```

Search completed: November 25, 2003, 12:09:32
Job time : 2555 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 10:20:53 ; Search time 74 Seconds
(without alignments)
1652.204 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGDECEPAPPEARPH.....HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 56978 seqs, 220691566 residues

Word size: 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US0989522/rnucat.21112003.184500.5896/app_query.fasta.1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n_ol1.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0989522.QCGN_1_1.56.0.rnucat.21112003.184500.5896 -NCPU=6 -ICPU=3
-NO_MMAP -LAQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/2/ina/Backfill1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	277	100.0	934	4	US-09-536-647-1
2	277	100.0	1288	4	US-09-620-312D-546

ALIGNMENTS

RESULT 1
US-09-536-647-1
; Sequence 1, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen

APPLICANT: Johnson, Randall
TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
FILE REFERENCE: GP50020
CURRENT APPLICATION NUMBER: US/09/536,647
CURRENT FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 834
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(780)
US-09-536-647-1

Alignment Scores:
Pred. No.: 1.75e-279 Length: 834
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-896-522-2 (1-277) x US-09-536-647-1 (1-834)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 1 ATGGCTTCGGCGGAGGCGCAACATCGCGAGCCCCCGCGCCGAGCCGATCGCGCAC 60

QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
DB 61 CAGCGGCGCTTCTGATAGAGGGGTAGCGCGGCGGCACTGCCAGCGGAGATCGACCTGTGT 120

QY 41 GlnValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
DB 121 GAGAAGATCATGAGTGTGCTGGACAGAAAGAGTGAACAGCGGAGCGGAGGAGGTGCTC 180

QY 61 IleLeuSerGlnAspArgPheTyrIleValIleThrAlaGluGlnLysAlaValAlaLeu 80
DB 181 ATCTCGAGCGCAGAGACAGGTTCTACAGGCTCTGACGCGCAGACCAAGGCGCTTG 240

QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 241 AAGGACAGTACATTTTGAACCATCAAGATGCTTTGACAAATGATTTGATGACAGAGCT 300

QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
DB 301 CTGAAGAACATGTGTGAGGCGCAAAACGGTGGAGGTCGACCTATGATTTGTGACACAC 360

QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValIleuPheGluGlyIle 140
DB 361 TCAAGTTTACAGAACACACGCGTGTCTACCTCGCGGACGTGTCTGTGTGAGGCGCATC 420

QY 141 LeuValPheTyrSerGlnGluIleLeuArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 421 TTGGTGTCTTACAGCAGAGATCCGGGACATGTTCCATCGCGCTCTCTGTGACACAC 480

QY 161 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyYrArgAspLeu 180
DB 481 GACTCGAGCTACAGGCTGTCTGAAAGATTTCCCGGAGCGTGGCGCGAGGAGGAGACCTG 540

QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 541 GAGCAGATTTCTGACCCAGACACACACCTTCGTAAGACCGGCTTCGAGAGATTCGCTCG 600

QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValIleAspAsnMetValAlaIle 220
DB 601 CCGACAAAGAAAGATATGCGAGTGTATCATCCACAGAGAGTGAATATATGATGTCCATC 660

QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleTyrPheArg 240
DB 661 AACCTGATGTGACACACATCCAGACATTTGGAATGTGTGACATGTGCAAAATGGACCGCA 720

```

QY      241  G|YGLVSeSaenGIYArgSerTYrLYsArgThrPheSerGIuProGIYAspHisProGIY 260
Db      721  GGAAGGCTCCATGGCGAGCTACAGCGACCTTTCTTGAGCCAGGAGCACACCTGGG 780

QY      261  MetLeuThrSerGIYLYsArgSerHisLeuGIuSerSerSerARgProHis 277
Db      781  ATGCTGACCTCTGGCAACGCTCACATTGGAGTCCAGACAGACACCCAC 831

RESULT 2
US-09-620-312D-546
; Sequence 546, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc FL_genes Version 1.0
; SEQ ID NO 546
; LENGTH: 1288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(928)
US-09-620-312D-546

Alignment Scores:
Pred. No.:      2,64e-279      Length:      1288
Score:          277.00         Matches:      277
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     100.00%       Indels:        0
DB:              4            Gaps:         0

US-09-896-522-2 (1-277) x US-09-620-312D-546 (1-1288)

QY      1  MetLAserAlaGIYGLuAspCYsGIuSerProAlaProGIYAlaAspArgProHis 20
Db      95  ATGGCTTCGGGGGAGGGAAGACTGCGAGAGCCCGCGCGAGCGACCGTCCGCAC 154

QY      21  GlnArgProPheLeuLeuIGLYValSerGIYGLYThrAlaSerGIYLYsSerThValCys 40
Db      155  CAGCGGCGCTTCCGTGATAGGGGTGAGCGCGCGCACTGCCAGCGGAGAGTGCACCGTGTGT 214

QY      41  GlnLYsIleMetGIuLeuLeuIGLYLnaenGIuValGIuGIuArgGIuArgLYsValVal 60
Db      215  GAGAAATCATATGAGATTGCTGGGACAGAAAGAGGTGAAACAGCGGACGGAAGTGTGTC 274

```

```

QY      61  IleLeuSerGIuAspArgPheTYrLYsValIleuThrAlaGIuGIuLYsAlaLYsAlaLeu 80
Db      275  ATCTTAGCCAGGACACGTTCTTACAGGTCTTACCGGACAGACGAAAGGCCACTTG 334

QY      81  LYsGIYGLnTYrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      335  AAGGACAGTACATTTTGGATCCATCCAGATGCCCTTGATATATGATTGATGACACAGACT 394

QY      101  LeuLYsAsnIleValIGLYsThrValIGLYuValProThrTYrAspPheValThrHis 120
Db      395  CTGAAGAACATCTGTGAGGGGCAAAACGGTGAAGTCCACCTATATTTGTGACACAC 454

QY      121  SerArgLeuProGIuThrThrValValTYrProAlaAspValIleuPheGIuGIYIle 140
Db      455  TCNAGGTTACAGAGACACAGCGGTGTCTTACCTCGCGAGCTGTTCTGTTAGGGCATC 514

QY      141  LeuValPheTYrSerGIuGIuIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      515  TTGGTGTCTTACAGCGAGGATCCGGGACATGTTCCACTGGGCTTCTCGTGACACC 574

QY      161  AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGLYArgAspLeu 180
Db      575  GACTCCGACGTGCGGTGTCTCGAAGAGTCTCCGGGAGGTGGCCGAGGAGGACCTG 634

QY      181  GIuGIuIleLeuThrGIuTYrThrThrPheValLYsProAlaPheGIuGIuPheCysLeu 200
Db      635  GACCAATTTCTGACGACGATCACCACTTCTGTGAAACCGGCTTCCAGAGATTCTGCTG 694

QY      201  ProThrLYsLYsTYrAlaAspValIleIleProArgGLYValAspAsnMetValAlaIle 220
Db      695  CCGACAAAGAAATGATCCGATGTGATCATCCGACGAGAGTGGAACAATATGGTTGCCATC 754

QY      221  AsnLeuIleValGIuIleGlnAspIleLeuAspIleLeuAsnGLYAspIleCysLYsTYrPhe 240
Db      755  AACCTGATCGTGCAGACATCCAGACATTTCTGAATGTGACATCTGCAAAATGGCACCGA 814

QY      241  G|YGLVSeSaenGIYArgSerTYrLYsArgThrPheSerGIuProGIYAspHisProGIY 260
Db      815  GAGGGGTCCATGGGGGAGCTTCAAGCGGACCTTTCTGAGCCAGGAGGACCACTTGGG 874

QY      261  MetLeuThrSerGIYLYsArgSerHisLeuGIuSerSerSerARgProHis 277
Db      875  ATGCTGACCTCTGGCAACGCTCACATTGGAGTCCAGACAGACCCAC 925

```

Search completed: November 25, 2003, 12:10:48
Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 08:18:06 ; Search time 318 Seconds
(without alignments)
2351.397 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277
Sequence: 1 MASAGEDESPAPFADRPH.....HPGMLTSGKSHLESSRRPH 277

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Word size: 15

Total number of hits satisfying chosen parameters: 30

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=xlh
-Q=/cgr2.1/USPTO.spool/US09896522/rnatc.21112003.184458.5860/app_query.facta.1.455
-DB=N.Geneseq.19Jun03 -QPMT=fstap -SUFFIX=p2n_ol1.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL
-OUTPM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09896522.@CGN.1.1.312.@runat.21112003.184458.5860 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: N.Geneseq.19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	277	100.0	834	23	AAK98735	DNA of a human uri
2	277	100.0	1288	22	AA158658	Human polynucleoti
3	277	100.0	1624	24	AA158658	Human 57658 DNA.
4	277	100.0	2152	22	AA160444	Human polynucleoti
5	277	100.0	2160	22	AA115621	Human cDNA sequenc
6	191	69.0	753	22	AA104832	Human cDNA clone (
7	189	68.2	900	22	AA175355	Human uridine kina
8	136	49.1	1396	24	AA189762	Human polynucleoti
9	112	40.4	734	24	AA151801	Human mdt cDNA in
10	60	21.7	9732	22	AA136439	Human musculoskele
11	60	21.7	9732	25	AA136439	CDNA encoding nove
12	60	21.7	19125	22	AA136440	Human musculoskele
13	60	21.7	19125	25	AA136440	CDNA encoding nove
14	25	9.0	715	20	AA121545	Human gene expres
15	25	9.0	1310	21	AA115785	Human prostate can
16	25	9.0	1322	22	AA123801	Human transferrase
17	25	9.0	1322	24	AA157850	Human ovarian anti
18	25	9.0	1402	22	AA157850	Human polynucleoti
19	25	9.0	3881	23	ABV23998	Human prostate exp
20	25	9.0	3881	23	ABV24102	Human prostate exp
21	25	9.0	3881	23	ABV24281	Human prostate exp
22	25	9.0	3881	23	ABV29881	Human prostate exp
23	25	9.0	3881	23	ABV29983	Human prostate exp
24	25	9.0	3881	23	ABV30091	Human prostate exp
25	15	5.4	65	24	ABN27650	Rat spliced trans
26	15	5.4	65	24	ABN51411	Mouse spliced tran
27	15	5.4	549	25	ABT33606	Anticancer agent t
28	15	5.4	822	23	AB110981	Drosophila melanog
29	15	5.4	1116	22	AA159636	Human polynucleoti
30	15	5.4	3413	23	AB110980	Drosophila melanog

ALIGNMENTS

RESULT 1	AAK98735	standard; DNA; 834 BP.
ID	AAK98735	
XX	AAK98735	
AC	AAK98735	
XX	AAK98735	
DT	02-MAY-2002	(first entry)
XX	02-MAY-2002	
DE	DNA of a human uridine kinase (UDK).	
XX	DNA of a human uridine kinase (UDK).	
KW	Human; uridine kinase; diagnostic assay; mutation detection; UDK;	
KW	probe; chromosome localisation study; tissue expression; gene therapy;	
KW	antibody; vaccine; human ovarian cancer; immunological disorder;	
KW	human colon carcinoma; immunogen; ds.	
XX	human colon carcinoma; immunogen; ds.	
OS	Homo sapiens.	
XX	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1..780
FT		/tag= a
FT		/partial
FT		/note="No stop codon"
XX		
PN	WO200172963-A2.	
XX	WO200172963-A2.	
PD	04-OCT-2001.	
XX	04-OCT-2001.	
PF	27-MAR-2001; 2001WO-US09663.	
XX	27-MAR-2001; 2001WO-US09663.	
PR	27-MAR-2000; 2000US-0536647.	
XX	27-MAR-2000; 2000US-0536647.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.	

PI Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
DR P-PSDB; AAOI4412.
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PR treating conditions associated with uridine kinase imbalance -
XX
XX Claim 5; Page 22-23; 31pp; English.

The invention relates to newly identified human uridine kinase (UDK)
polypeptides and polynucleotides and methods for producing such
polypeptides by recombinant techniques. Also disclosed in the invention
are methods for utilising uridine kinase polypeptides and polynucleotides
in diagnostic assays. The polynucleotides and polypeptides of the
invention may be used as diagnostic reagents by detecting mutations in an
associated gene. An array of oligonucleotide probes comprising the
uridine kinase polynucleotide sequence or fragments thereof can be
constructed to conduct efficient screening of genetic mutations, for
example. Detection of abnormally decreased or increased levels of
polypeptide or mRNA expression may also be used for diagnosing or
determining susceptibility of a subject to a disease of the invention.
The polynucleotide sequences of the invention can be used for chromosome
localisation studies and tissue expression studies. The polypeptides of
the invention or fragments thereof may be used as immunogens to produce
antibodies. These antibodies may be employed to isolate or identify
clones expressing the polypeptide. The polypeptides and polynucleotides
of the invention can be used as a vaccine or in gene therapy to treat
diseases such as human ovarian cancer, human colon carcinomas, and
immunological disorders. This polynucleotide sequence represents the DNA
of a human uridine kinase of the invention.

Sequence 834 BP; 194 A; 219 C; 265 G; 156 T; 0 other;

Alignment Scores:

Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

2e-269	length:	834
277.00	Matches:	277
100.00%	Conservative:	0
100.00%	Mismatches:	0
100.00%	Indels:	0
23	Gaps:	0

US-09-896-522-2 (1-277) x AAK98735 (1-834)

Dy
1 MetAlaSerAlaGlyGlyGlnaspCyseGluSerProAlaProGluAlaAspArgProHis 20

Db
1 ATGGCTTCGGCGAGGCAGAAGACTGCCAAGACCCCCCGCCGGAGCCCCCAGCCGTCCGCA 60

QY 21 GlnArgProPheLeuIleGlyValSerCylgYlThrAlaSerGlyLysSerThrValCys 40
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 CAGCGCCCTTCCTGATAGGGGTGACGGCGCACTGCAGCGGAAGTTCGACCGTGCT 120

Qy 41 G l u y s i l e m e t G l u e u l e u c i y g i n a s n g l u V a l g l u g l m a r g i n a r g l y s V a l V a l 60
 Db 121 G A G A A G A T C A T G A G A T T C T G G S A C A G A C G A G G T G A A C A G C G G C A G C G A A G G T G T C 180

QY ILELEISERGIMASPARGPHEYYRGLYSALAEUTHRALAGLUSINLYSALAVALAEN 80
DB 181 ATCTTGAGCCGAGGACAGSTTCTTACAAGGTCTGACGGCAGACAGAAGGCCAACGCCTTG 240

81 LysGlyGlnIlyTraspneaspIsIPCoBpaIapneaspasIspleuethsargInr 100
 |||||
 Db 241 AAAGGACAGTCAATTGGACCATCCGATGCTTTCGACAAATGATTGATGCACAGACT 300

Db 301 CTGAAGACATCTGGAGGGCAAAACGCTGGAGCTGCCGACCTATGATTTTGACACAC 360

QY 121 SerArgLeuProGluThrThrValValTyrProAlaaspValValLeuPheGluGlyLe 146
|||
DB 361 TCAGGTTACAGAGACCACGGTGTTCACCTCGGACGTGGTTCTGTTTGAAGGCATC 422

QY	141	LeuVal.PheTyr.SerGlnIuIleArgPheMet.PheHisIleuArgLeuPheVal.AspThr	160
Db	421	TrpGtGtTCTACGCCAGAGAGATCCGGACATGTTCCACTCCGCTCTTCGTGGACACC	488
QY	161	AspSerAspValArgIleuSerArgArgValIleuArgAspValArgArgIArgAspLeu	180
Db	481	GACTCCGAGGTGAGGTGTCCTGAGAGAGTTCTCCGGACGTCGCCCGAGGGAGGACTG	540
QY	181	GluGlnIleLeuThrGlnTyrThr.PheVal.LysProAlaPheGluGluPheCysLeu	200
Db	541	GACACAGATCTGACGAGATCACCACTTCGGAAAGCGGCTTCGAGAGATTCGTGCTG	600
QY	201	ProThrIleLysTyrValAspValIleIleProArgIValAspAsnMetValAlaIle	220
Db	601	CCAGCAAGAGATGATGCCATGTGATCATCCACAGAGAGTGGACAAATATGTTGCCATC	660
QY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg	240
Db	661	AACCTGATGCTGAGACATCCAGGACATTTGAAATGTGACATCTGCAATGTGACCGA	720
QY	241	GlyGlySerAsnGlyArgSerTyrIlySarGTrPheSerGluProGlyAspHisProGly	260
Db	721	GAGGGGGCCAAATGGGGGAGCTACAAAGGGAGCACTTTTGTGACCGAGGGACACACCTGGG	780
QY	261	MetIleuThrSerGlyLysArgSerHisIleuGluSerSerSerArgTrpProHis	277
Db	781	ATGCTGACTCTGGCAAAAGGTCACTATTGGAGTCCAGCAGACAGACCCAC	831

ID AAI58658 standard; cDNA; 1288 BP.

AC AAI586

DT 22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 861.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

Alzheimer's disease; Parkinson's disease; Huntington's disease; haemostatic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

28 XX
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860

XX
W00001E3310-21XX 36-TIT-2001
PD

XX 36-DEC-2000: 2000WD-TT934263

XX 21-JAN-2000: 2000ITS-0488725
PR

PR 25-APR-2000; 2000US-0552317
PR 09-JUL-2000; 2000US-0598042

PR	19-JUL-2000; 2000US-062031Z
PR	03-AUG-2000; 2000US-0653450

PR 14-SEP-2000; 2000US-0662191
PR 19-OCT-2000; 2000US-0693036

XX 29-NOV-2000; 200005-0121344

XX (HISE-1 HISEL INC.
XX

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX
E1 2100 25, 2000 10, 2100 10
E1 2100 25, 2000 10, 2100 10

DR P-PSDB; AAM39502.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 861; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA1642213) with nootropic,
 CC immunosuppressant and cytoprotective activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 1288 BP; 275 A; 361 C; 418 G; 234 T; 0 other;

Alignment Scores:

Pred. No.: 3e-269 Length: 1288
 Score: 277.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-896-522-2 (1-277) x AA158658 (1-1288)

QY 1 MetAlSerAlaGlyGluAspGlySerProAlaProGluAlaAspArgProHis 20
 DB 95 ATGGCTTCGGCGGAGGCGGAAAGCTCGAGAGCCCGCGGAGGCGCGCTCCGAC 154
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 155 CAGCGGCCCTTCCTGTAAGGAGGAGGCGGAGCTGCGCAGCGGAAAGTGCACCGTGTG 214
 QY 41 GlnValIleMetGlnLeuLeuGlyGlnAsnGluValGlnGlnArgGlnArgValVal 60
 DB 215 GAGAAATCATGATGAGTTCCTCGGACGAGACGAGTGGAAACGCGGACCGGAGGTGTC 214
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaValAlaLeu 80
 DB 275 ATCTGAGCGCAGGACGAGTCTCAAGGTCCTGACGCGCAGACGAGGCGCAAGCCCTTG 334
 QY 81 LysGlyGlnTyrTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 335 AAGGACAGTACAAATTTTGAACATCCAGATGCCCTTGATATATATATTGATGACAGGACT 394
 QY 101 LeuLysAsnIleValGlnGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 395 CTGAAATAATCTGAGGCGCAAAACGCTGAGAGTCCGACCTATGATTTTGGACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
 DB 455 TCMAAGTTTACAGAGACCAACGCTGCTACCCCTGCGGACGTGTTCTGGAGGCGATC 514
 QY 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisIleLeuArgLeuPheValAspThr 160
 DB 515 TTGGTTTCTACACGAGGAGATCCGAGCATTTTCCACTGTGGCTCTTCGTTGGACACC 574
 QY 161 AspSerAspValArgLeuSerArgValValLeuArgAspValValArgArgGlyArgAspLeu 180
 DB 575 GATTCGAGCTCAGGCTGTCTCGAAGATTTCTCCGGGACCTGCGCGGAGGAGGACCTG 634
 QY 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGluPheCysLeu 200
 DB 635 GAGCAGATTCTGACGAGTACACACACCTTCGTGAACCGGCTTCGAGGAGTCTGCTG 694

QY 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 DB 695 CCGACAAAGAAGATATGCCATGTGATCTCCACAGAGAGTGGACAAATATGTTGGCATC 754
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
 DB 755 AACCTGATGTCAGACATCCAGACATTTCTGAATGTGATGATCTGCAAAATGGCACCGA 814
 QY 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGlnProGlyLysAspHisProGly 260
 DB 815 GAGGGGTCCAAATGGGCGGAGCTTCAACGACCTTTTCTGAGCCAGGAGGAGCACTCTGG 874
 QY 261 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerSerArgProHis 277
 DB 875 ATGCTGACCTCTGGCAACGTCATCTTTGGAGTCCAGACGACACCCAC 925
 RESULT 3
 AAD27186
 ID AAD27186 standard; DNA; 1624 BP.
 XX
 AC AAD27186;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 DNA.
 XX
 KW Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KW 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KW rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KW sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW gene therapy; chromosome mapping; tissue typing; dermatological;
 KW cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KW anticonvulsant; db.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 94..927
 FT /tag= a
 FT /product= "Human 57658 protein"
 FT misc_feature 94..924
 FT /tag= b
 FT /note= "This region is specifically claimed as
 SEQ ID NO: 3 in claim 1 of the specification"
 XX
 PN WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 DR WPI; 2002-140091/18.
 DR P-PSDB; AA161592.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 PS Claim 1; Fig 1a; 103bp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,

CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology), and predictive medicine (e.g. diagnostic assays).
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is a DNA encoding human 57658 protein.

XX
 SQ Sequence 1624 BP; 354 A; 427 C; 506 G; 337 T; 0 other;

Alignment Scores:

Pred. No.:	3,72e-269	Length:	1624
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-896-522-2 (1-277) x AAD27186 (1-1624)

QY 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 DB 94 ATGGCTTCGGCGGAGGAGGAGCTGGAGAGCCCGCGCGGAGCCGACCGTCCGCAC 153
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 DB 154 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGACCTGCGACGGGAGAGTGCACCGTGTGT 213
 QY 41 GlnuylleMetGluLeuLeuGlyGluAsnGluValGluGlnArgGlnAsnGlyValVal 60
 DB 214 GAGAAAGATCAGAGGATGCTGGGACAGAAACGAGGTGGAAACGCGCAGCGAAAGTGTCTC 273
 QY 61 IleuSerGlnAspArgPheTyrlsValLeuThrAlaGluGlnLysAlaValAlaLeu 80
 DB 274 ATCTGAGCCGAGGACAGGTTTCAAGGTCTTGACGGCAGAGCGAAAGGCCAGGCTTGG 333
 QY 81 LysGlyGlnTyrlsAsnAspPheAspPheAspAlaPheAspAsnAspMetHisArgThr 100
 DB 334 AAAGAGACGATCAATTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACT 393
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrlsAspPheValThHis 120
 DB 394 CTGAAGAACATCGTGGAGGCGAAACGCTGGAGGTGCGACCTATGATTTGTGACACAC 453
 QY 121 SerArgLeuProGluThrThrValValTyrlsProAlaAspValValLeuPheGluGlyIle 140
 DB 454 TCAAGGTTTCCAGAGACCAAGGTGTCTACCTCCGAGCGTGTCTGTTTGAAGGCATC 513
 QY 141 LeuValPheTyrlsSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 DB 514 TTGGTGTTCACACGAGAGATCCGGGACATGTTCCACCTGCGCCTCTTGATGACACC 573
 QY 161 AspSerAspValAlaGluSerArgArgValLeuArgAspValAlaGluGlyArgAspLeu 180
 DB 574 GACTCCGACGTCAGGCTGTCTCGAAGGTTCTCCGGACGCGCCGACGAGGAGGACCTG 633
 QY 181 GlnGlnIleLeuThrGlnTyrlsThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 634 GAGAGATTCAGAGCGCATACCACTTCTGTAAGCCGCGCTTCGAGAGTTTGTGCTG 693
 QY 201 ProThrLysLysTyrlsAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
 DB 694 CCGCAAAAGAGTATGCGATGTATCATCCAGAGAGTGAGCAATATATGTTGCCATC 753
 QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
 DB 754 AACCTGATCGGAGCAGCATCCAGACATTTTGATAGGTGACATCTCAATATGACACCA 813
 QY 241 GlyGlySerAsnGlyArgSerTyrlsArgThrPheSerGluProGluAspHisProGly 260

DB 814 GAGAGTTCACATGGCGGAGCTACAGCGGACCTTTTTCAGAGCAGAGGAGACCACTGGG 873

QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 DB 874 ATGCTGACCTCTGGCAACGGTTCACATTTGAGTTCAGAGCAGACAGACCCGCAC 924

RESULT 4

AA160444
 ID AA160444 standard; cDNA, 2152 BP.

XX AA160444;

DE 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 4433.

DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.

XX Homo sapiens.

PN NO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US34263.

PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.
 DR P-PDB; AAM41288.

PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 4433; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actvatin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 2152 BP; 468 A; 581 C; 660 G; 442 T; 1 other;

Alignment Scores:

Pred. No.:	4,84e-269	Length:	2152
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-2 (1-277) x AAH160444 (1-2152)

```

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 59 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGCGAGGCGCGACCGTCCGCAC 118
QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThraIaSerGlyLysSerThrValCys 40
DB 119 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGGCACTGCCAGCGGGAAGTCGACCGTGTGT 178
QY 41 GlnLysIleMetGluLeuGluGlnAsnGluValGluGlnArgGlnArgGlyValVal 60
DB 179 GAGAGATTCATGGAGTTCGTGGGACAGACGAGGTGGAACGCGGACGCGAGAGTGTCTC 238
QY 61 IleuSerGlnAspArgPheTyTyLeValIleuThrAlaGluGlnLysAlaLysAlaLeu 80
DB 239 ATCTGAGCCAGGACGAGGCTTCAAGAGTCTCTGACGCGACGAGCGAGGCGCAAGGCTTG 298
QY 81 LysGlyGlnTyTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
DB 299 AAGGACAGATACATTTTGAACATCCAGATGCCCTTATATATGATTTGATCCACGAGACT 358
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyTrpAspPheValThrHis 120
DB 359 CTGAGAGACATCTCGGAGGCGAAGACGTTGAGGTGCCGACTTATTTTGACACAC 418
QY 121 SerArgLeuProGluThrThrValValTyTrpAlaAspValIleuPheGluGlyIle 140
DB 419 TCAAGGTTACAGAGACACACGTTGTCTACCTCGGACGCGTCTGTTGAGGCGATC 478
QY 141 LeuValPheTyTrsGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 479 TTGGTGTCTTACAGCCAGATCCGGGACATGTTCCACCTGCCCTCTTCGTGGACACC 538
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValAlaArgGlyArgAspLeu 180
DB 539 GACTCCAGGTCAGGCTGTCTCGAAGATTCTCCGGACGTCGCCGAGGAGGAGACTTG 598
QY 181 GlnGlnIleLeuThrGlnTyTrpThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 599 GAGCAGATTCTGACGCGATCACCACTTGTGTAAAGCCGCTTCGAGGAGTTCGCTG 658
QY 201 ProThrLysLysTyTrpAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
DB 659 CCCACCAAGAGATATGCCGATGATATCCCGCAGAGATGAGCAATATGTTGTCATC 718
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
DB 719 AACCTGATGTCGAGCAGCATTCAGACATTCGTGATGTACATCTGCAATATGCGACCGA 778
QY 241 GlyIYSerAsnGlyArgSerTyTyLeArgThrPheSerGluProGlyAlaAspHisProGly 260
DB 779 GAGAGGTCGAATGGCGGAGCTACAAAGCGACCTTTCTGAGCAGGAGGACACACCTGGG 838
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
DB 839 ATGCTGACCTCTGGCAACGATCATTGTGAGTCCAGAGCAGACCCGCAC 889

```

RESULT 5

AAH15621
ID AAH15621 standard; cDNA; 2160 BP.
AC AAH15621;
XX

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99UP-0248036.

XX 27-AUG-1999; 99UP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;

XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any special methods. AAH0166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

SQ Sequence 2160 BP; 457 A; 591 C; 671 G; 441 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13951.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99UP-0248036.

XX 27-AUG-1999; 99UP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;

XX Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX PS Claim 8; SEQ ID 13951; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any special methods. AAH0166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

US-09-896-522-2 (1-277) x AAH15621 (1-2160)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 95 ATGGCTTCGGCGGAGGCGAAGACTGCGAGAGCCCGCGAGGCGCGACCGTCCGCAC 154


```

QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGGGGCCCTTCCTGATAGGGGTGAGCGCGCACTGCCACGGGAAGTCCAGCGTGTGT 214
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db 215 GAGAGAGATCATGAGATTCTGCGACAGAACGAGGTGGAACCGCGCAGCGAGAGGTGTC 274
QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGluIleValAlaLeu 80
Db 275 ATCTGAGCCAGCAGAGGTTCTTCAAGGTCCTGACGGACAGACAGAGCCAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspPheProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGACAGTACAATTGTTGACCATCCAGATGCTTTGATATGATTGATTCACAGAGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 395 CTGAGAACATCGTGGAGGCAAAACGTTGAGGTGCGACCTATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIle 140
Db 455 TCAAGGTTACCAAGACCAACGATGCTTACCTCGGACGTGTCTGTGTGAGGGCACTC 514
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 515 TTGGTGTCTACAGCAGAGATCCGGGACATGTTCCACCTCGCGCTCTTGTGTGACACC 574
QY 161 AspSerAspValAlaGluSerThrArgValLeuArgAspValAlaGArgGlyArgAspLeu 180
Db 575 GACTCCGACGTGAGGCTGTCTCGAAGAGTTCCTCGGAGCGTGGCCGAGGAGGACCTG 634
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
Db 635 GAGCAGATTCTGACGCACTACCACTTCGTGAAGCGCGCTTGAAGAGTCTCGCTG 694
QY 201 ProThrLysLysTyrAlaAspValIleLeuProArgGlyValAlaAspMetValAlaIle 220
Db 695 CCGCAAAAGAGTATGCCGATGTGATCATCCACAGAGAGTGGCAATATGTTGCCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrHisArg 240
Db 755 AACCTGATCGTGCAGCACATCCAGACATTTGATGATGTGACATCTGCMAATGCGACCGA 814
QY 241 GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGluAspHisProGly 260
Db 815 GGAAGGTCATAGGCGGAGCTTACAGCGACCTTTTCTGAGCCAGGGGAGCACCCCTGGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 277
Db 875 ATGCTGACCTCTGGCAACGCTCACATTGTGAGTCCAGACAGACCCCCAC 925

```

RESULT 6

AAH04832 standard; cDNA: 753 BP.

AAH04832;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1667.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

BP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

```

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
PI Ichii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T,
DR WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 1667; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 753 BP; 158 A; 196 C; 261 G; 135 T; 3 other;
SQ

```

Alignment Scores:

Pred. No.:	8,21e-183	Length:	753
Score:	191.00	Matches:	191
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.95%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-2 (1-277) x AAH04832 (1-753)

```

QY 1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 95 ATGGCTTCGGCGGAGCGGAAGACTCGAGAGCGCCCGCGGAGGCCAGCGTCCGCAC 154
QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db 155 CAGCGGCCCTTCTCTGATAGGGGTGAGCGCGGCACTCCAGCGGGAAGTCCGCGTGTG 214
QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValIle 60
Db 215 GAGAGAGATCATGAGTTGCTGGGACAGAACGAGTGAACAGCGGACGGAAGGTGTGC 274
QY 11eLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
Db 275 ATCTGAGCCAGCAGAGTTCCTGACAGGTCCTGACGCGAGAGAGAGCCAAAGGCTTG 334
QY 81 LysGlyGlnTyrAsnPheAspPheProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 335 AAAGACAGTACAATTGTTGACCATCCAGATGCTTTGATATGATTGATGTGACAGAGACT 394

```


QY 101 LeuValAsnIleValGluGlyValThrValGluValProThrTyrAspPheValThrHis 120
 DB 395 CTGAGAACATCTGTTGAGGGGCAAAACGTTGAGGTCGACCTATGATTTTGTACACAC 454
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
 DB 455 TCAAGGTACAGAGACACGCGTGTCTACCCCTCGGACGTCGTTCTGTTGAGGCATC 514
 QY 141 LeuValPheTyrSerGluGluIleArgAspMetPheHisLeuValGluPheValAspThr 160
 DB 515 TTGGTGTCTACACGCGAGATCCGGACATGTTCCACCTGCGCTCTTCGTGACACC 574
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGlyArgAspLeu 180
 DB 575 GACTCCACGTCAGCGCTGCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGGAGACCTG 634
 QY 181 GluGlnIleLeuThrGlnTyrThrThrPheVal 191
 DB 635 GAGCAGATTCTGACGACGATACACCACTTCGTCG 667

RESULT 7

AAH75355 ID AAH75355 standard; cDNA; 900 BP.

AC AAH75355;

DT 02-OCT-2001 (first entry)

DE Human uridine kinase encoding cDNA.

XX Human, uridine kinase; UK; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 8..838

FT /product= "uridine kinase"

FT /note= "Claimed in claim 1"

XX CN1287172-A.

XX 14-MAR-2001.

XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.

XX (UYFU-) UNIV FUDAN.

XX Yu L, Zhao Y, Zhang H;

XX WPI; 2001-409529/44.

XX P-PSDB; AAG64506.

XX Human uridine kinase and its coding sequence, preparation and

XX application -

XX Claim 1; Page 15(Disclosure); 20pp; Chinese.

XX The invention relates to human uridine kinase (UK).

XX Sequence 900 BP; 201 A; 237 C; 292 G; 170 T; 0 other;

Alignment Scores:
 Pred. No.: 1e-180
 Score: 189.00
 Percent Similarity: 98.47%
 Best Local Similarity: 98.47%
 Query Match: 68.23%
 DB: 22

US-09-896-522-2 (1-277) x AAH75355 (1-900)

QY 18 ArgProHisGlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSer 37
 DB 56 CGTCCGACACGAGGTCCTTCTGATAGGGGTAGGGCGGACCTGCGAGGAAAGTCG 115
 QY 38 ThrValCysGluIleMetGluLeuGluGlyValAsnGluValGluGlnArgGlnArg 57
 DB 116 ACCGTGTGAGAAAGATCATGAGATTGCTGGGACAGAACAGAGTGCAGACCGGCA 174
 QY 58 -LysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAl 77
 DB 175 TAAAGTGTATCTGAGCCAGAGAGGTCCTACAGATCTCTGACGCGAGAGAGAGGC 234
 QY 77 alyValAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 97
 DB 235 CAAGGCCCTTGAAGGACAGTAAATTTTGCATTCACATGCTTGTATATATATATGAT 294
 QY 97 tHisArgThrLeuValAsnIleValGluGlyValThrValGluValProThrTyrAspPh 117
 DB 295 GCAACGACTCTTAAGAACATCTGAGAGGCAAAACGTTGAGGTCGACCTATGATTT 354
 QY 117 eValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPh 137
 DB 355 TGTGACACACTCAAGGTTACAGAGACCAAGGTCGTCACCTGCGGACGTCGTTCTGTT 414
 QY 137 eGluGlyIleLeuValPheTyrSerGluGluIleArgAspMetPheHisLeuArgLeuPh 157
 DB 415 TGAAGGCATCTTGTGTCTACAGCCAGAGAGATCCGGACATGTTCCACTGCGCTCTT 474
 QY 157 eValAspThrAspSerAspValArgLeuSerArgValArgValLeuArgAspValArgArg 177
 DB 475 CGTGCACACCGACTCCGAGCTCAGGCTGTCTCAAAAGTTCTCCGAGCGTCGCGCGG 534
 QY 177 YArgAspLeuGluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGlu 197
 DB 535 GAGGAGACCTGGACAGATTTCTGACGACATACACCTTGTAAGCCGCGCTTCAGAGA 594
 QY 197 uPheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMe 217
 DB 595 GTTCTGCTGCGCCACAAAGAGATGCGATGATCATCCACGAGGTGACAAATAT 654
 QY 217 tValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
 DB 655 GGTTCATATACCTATCTGACGACATCCAGACATCTGAAAGGTGACATCTGCAA 714
 QY 237 sTrpHisArgGlyGlySerAsnGlyArgSer-TyrLysArgThrPheSerGluProGlyA 257
 DB 715 ATGGACCGAGAGAGGATCCATATGGCGGAC-GTACAAAGCGGACCTTTCTGACCGAGGG 773
 QY 257 sPheAspProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProH 277
 DB 774 ACCACCTCGGATGTGACTCTGCGCAAAACGTCACATTTGGATCGACAGACGACCCC 833
 QY 277 ts 277
 DB 834 AC 835

RESULT 8

ABL89762 ID ABL89762 standard; cDNA; 1396 BP.

AC ABL89762;

DT 24-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 324.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 XX valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 XX cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; human; secreted protein; gene; ss.

XX OS Homo sapiens.
 XX PN WO200190304-A2.
 XX PD 29-NOV-2001.
 XX PF 18-MAY-2001; 2001WO-US16450.
 XX PR 19-MAY-2000; 2000US-205515P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Birse CE, Rosen CA;
 XX DR WPI: 2002-122018/16.
 XX DR P-PSDB; ABB89353.
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders -
 XX Claim 4; SEQ ID NO 324; 2081bp + Sequence Listing; English.
 XX The invention relates to novel genes (ABU89449-ABU90853) and proteins (ABB89044-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 1396 BP; 317 A; 358 C; 425 G; 287 T; 9 other;
 Alignment Scores:
 Pred. No.: 3,84e-127 Length: 1396
 Score: 136.00 Matches: 244
 Percent Similarity: 97.99% Conservative: 0
 Best Local Similarity: 97.99% Mismatches: 3
 Query Match: 49.10% Indels: 5
 DB: 24 Gaps: 0
 US-09-896-522-2 (1-277) x ABU89762 (1-1396)
 QY 1 Meta1aSerAlaGlyGlyAlaSPCySerProAlaProGluAlaAspArgProHis 20
 Db 41 ATGGCTTGGCGGCGAGCGAGACTGCGAGAGCCCGCGGAGCGCCGACCTCCGAC 100
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThra1a-SerGlyysSerThrValCy 40
 Db 101 CAGCGGCCCTCTCGATAGGGGTGAGCGCGGCACTG-CAGCGGGAAGTGCAGCGTGTG 159
 QY 40 sGluValIleMetGluLeuLeuGlyGlnArgGluValGlnGlnArgGlnArgGlyValVa 60
 Db 160 TGAAGAAGTCAATGAGTGTGCGGACAGAAAGAGTGTGACGCGCGAGCGAAGTGTGT 219
 QY 111eUeSerGlnAspArgPheTyrLysValIleThra1aGlnGlnIlysa1aLysAlaLe 80
 Db 220 CATCTTAGCCAGAGCAGGTTCTTACAGGCTCTTACGCGCAGACAGAGAGGCTT 279
 QY 80 uLySGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgTh 100

Db 280 GAAAGCAGCTACCAATTTTACCATCCAGATGCTTTGATTAATGATTGTGACAGAGAC 339
 QY 100 rLeuLysAsnIleValGlnGlySerThrValGluValProThrTyrAspPheValThrH1 120
 Db 340 TCTGAAGAACATCGTGAAGGGCAAAACGGTGAAGGGTCCAGCTTAATGATTTTGTACACA 399
 QY 120 sSerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyL1 140
 Db 400 CTCAGGTTACAGAGACCCAGGTGTCTACCTTCGCGAGCTGTCTGTTAAGGGCAT 459
 QY 140 eLeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspTh 160
 Db 460 CTGTGTCTTACAGCAGCAGAGATCCGGGACATGTTCCACCTGCGCTTCTGTGACAC 519
 QY 160 rAspSerAspValArgLeuSerArgArgVal-LeuArgAspValArgArgGlyArgAspL 180
 Db 520 CGACTCCGACGATGAGGTGTCTGAAAGAKT-TCTCCGGACGTGCGCGGAGGAGGACC 578
 QY 180 eugGlnIleLeuThrGlnTyrThrThrPheValIlyspProAlaPheGluGluPheCysL 200
 Db 579 TGAAGCAGATTCTGACGCA-TACACACCTTCTGTGAAGCGGCTTCGAGGATTTCTGCC 637
 QY 200 euProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaI 220
 Db 638 TGCAGCAAAAGATATGCCGATGTGATCATCCACGAGAGAGTGACAAATATGTTGCCA 697
 QY 220 leuLeuIleValGlnHisIleGlnAspIleLeuLeuGlnArgPheCysLysTyrHisA 240
 Db 698 TCAACCTGATCGGACGACATCCAGACATTCTGAATGGTGCATCTGCAATGGGCACC 757
 QY 240 rGlyGlySerAsnGlyArgSer 247
 Db 758 GAGGAGGTCCAATGGCGGAGT 760
 RESULT 9
 ABS51801
 ID ABS51801 standard; cDNA; 734 BP.
 XX
 AC ABS51801;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human mddt cDNA Incyte ID No: LI:235557.12:2001JAN12.
 XX
 KW Human; molecule for disease detection and treatment; MDDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antiasthmatic; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 XX WO200255738-A2.
 PN 18-JUL-2002.
 PD 09-JAN-2002; 2002WO-US01008.
 PF
 XX
 XX 12-JAN-2001; 2001US-261622P.
 PR 16-JAN-2001; 2001US-261865P.
 PR 17-JAN-2001; 2001US-262208P.
 PR 17-JAN-2001; 2001US-262209P.
 PR 17-JAN-2001; 2001US-262326P.
 PR 19-JAN-2001; 2001US-263063P.
 PR 19-JAN-2001; 2001US-263065P.
 PR 19-JAN-2001; 2001US-263329P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Panzer SR, Lincoln SR, Altus CM, Dufour GE, Hillman JL, Jones AL,
 PI Dam TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen AD;
 PI Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;

XX WP1: 2002-590679/63.
DR P-PSDB; ABG70328.
XX
PT New disease detection and treatment molecule (MDT) polynucleotides and
PT polypeptides, useful in diagnosing, studying, preventing or treating
PT diseases associated with MDT expression, e.g. autoimmune or
PT inflammatory disorders
XX
XX Claim 1; Page 98; 129pp; English.
XX
CC The present invention relates to the isolation of novel human
CC molecules for disease detection and treatment (MDT), and the
CC polynucleotide sequences (mdt) encoding them. The MDT polypeptides
CC may be used to screen for molecules that bind to, or are bound by the
CC encoded polypeptides, and to develop a transcript image of a tissue or
CC cell type. Probes comprising at least 20 nucleotides of the mdt
CC polynucleotide may be used to assess the toxicity of a test compound.
CC The MDT polypeptides and mdt polynucleotides are useful in the
CC diagnosis, study, prevention and treatment of diseases associated with
CC the expression of molecules for disease detection and treatment. Such
CC disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC cirrhosis, or cancers), and autoimmune/inflammatory disorders
CC (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdt
CC polynucleotides may also be used as molecule markers, in microarrays,
CC and in somatic or germline gene therapy. ABS51779-ABS51814 encode
CC the MDT proteins of the invention.
XX
SQ Sequence 734 BP; 166 A; 181 C; 226 G; 154 T; 7 other;

Alignment Scores:
Pred. No.: 3 21e-103 Length: 734
Score: 112.00 Matches: 112
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.43% Indels: 0
Gaps: 0
DB: 24

US-09-896-522-2 (1-277) x ABS51801 (1-734)

QY 90 AspAlaPheAspAspLeuMetHisArgThrLeuLysAsnIleValGluGlyLysThr 109
Db 14 GATGCCCTTGATATGATTGATGACAGGACTCTGAAGAATCGTGGAGGCAAAACG 73

QY 110 ValGluValProThrTyrAspPheValThiSerArgLeuProGluThrValVal 129
Db 74 GTGAGGAGTCCGACCTATGATTTGTGACACATCAAGATTACAGAGACCAACGTCGTC 133

QY 130 TyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGluIleArg 149
Db 134 TACCCCTCGGACGCGTTCCTTTGAGGGCATCTTGATGTTCTACACCGAGATCCGG 193

QY 150 AspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
Db 194 GACATGTTCCACGCGCCCTCTTGTCGACACGACCTCCGACGTCGTCGAGAAGA 253

QY 170 ValLeuArgAspValArgArgGlyArgAspLeuGluIleLeuThrGlnTyrThrThr 189
Db 254 GTTCTCGGAGACGCGCCGACGAGGAGGACCTGAGCAGATTTGACGACGATACACACC 313

QY 190 PheValLysProAlaPheGluGluPheCysLeuPro 201
Db 314 TTCGTGAGCCGCGCTTCGAGAGATTTCTGCTGCCG 349

RESULT 10
AAL36439
ID AAL36439 standard; DNA; 9732 BP.
XX
AC AAL36439;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2804.

XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiasthmatic; hepatotropic; antidiabetic; antiinflammatory; antiviral;
KW vintnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218390.
PR 26-JUL-2000; 2000US-0220863.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251856.
PR 06-DEC-2000; 2000US-0251856.
PR 06-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SW;
XX
DR WPI; 2001-451937/48.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
diagnosis -
PS
XX Example 2; SEQ ID NO 2804; 781pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (AAJ34669-AAJ37666) and proteins
CC (AAB01087-AAB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;

Alignment Scores:
Pred. No.: 8..93e-50 Length: 9732
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
Gaps: 0

US-09-896-522-2 (1-277) x AAJ36439 (1-9732)
Qy ValAlaIleasnLeuIleValjGlnHisIleGlnAspIleLeuAsnGIYAspIIeCYslYS 237
Db GTTCCATCAACTGTCTGTGGAGCACAICCGACATTCTAAATGGTCACATCTGCCAA 3966
Qy TrpHisArgGIlySerAnsnGIYArSerrTYLYrArGrThPhSeSGlUProGIYAsP 257
Db TGGCACCGAGAGAGGTCCAATGGCGGAGCTCAAGCGGACCCTTTTGAGGCAGAGGAC 4020
Qy HisProGIYMeIerLeuThSerGIYArSArSrHisIleuGlUsErSerSArGrProHis 277
Db CACCCTGGAGTGTGAACCTCTGGCAAAACGGTCACATTGGATGCCAGCAGACCCAC 4086

RESULT 11
ID ABX59427 standard; cDNA; 9732 BP.
XX AC ABX59427;
XX DT 26-FEB-2003 (first entry)
XX DE cDNA encoding novel human musculoskeletal system antigen #1771.

XX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
 KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KM post-operative tissue repair; limb regeneration; neuronal growth;
 KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KM AIDS-related complex; chondrocyte growth; bone regeneration;
 KM periodontal regeneration; tissue transport; bone graft; skin aging;
 KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KM cell growth; organ transplant; cell differentiation; body height;
 KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KM pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
 KM depression; tendency for violence; pain; reproductive capability;
 KM hormone level; endocrine level; appetite; libido; memory; stress;
 KM storage capability; fat content; lipid content; protein content;
 KM carbohydrate content; vitamin content; cofactor content;
 KM nutritional component.
 XX Homo sapiens.
 XX US2002147140-A1.
 PN 10-OCT-2002.
 XX 17-JAN-2001; 2001US-0764877.
 PF 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-220964P.
 PR 14-AUG-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-229509P.
 PR 05-SEP-2000; 2000US-229513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234223P.
 PR 21-SEP-2000; 2000US-234227P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236807P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239335P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.

PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC,
 XX WPI; 2003-128199/12.
 DR
 XX
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 PT associated polypeptides, useful for detecting disorders, e.g. cancer -
 PS disclosure, SEQ ID NO 2804; 321pp, English.
 XX
 XX The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, haematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's mental state or physical state by influencing biorhythms;
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
 XX
 SQ Sequence 9732 BP; 2063 A; 2805 C; 2561 G; 2303 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,93e-50 Length: 9732
 Score: 60.00 Matches: 60
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 21,664 Indels: 0
 DB: 25 Gaps: 0
 US-09-896-522-2 (1-277) x ABX59427 (1-9732)
 QY 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
 Db 3901 GTTGGCATCAACTGATCTGTCGACGACATTCGATTCGATTCGACATTCGACAA 3960
 QY 238 TTPHISARGGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAsp 257
 Db 3961 TGGCACCGAGGAGGGCTCAATGGCGGAGCTCAAGACGGAGCCTTTTCTGAGCCAGGAGAC 4020

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-451937/48.
XX
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX
XX Example 2; SEQ ID NO 2805; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABR03087-ABR04109) associated with the musculoskeletal system useful
XX for preventing, treating or ameliorating medical conditions e.g. by
XX protein or gene therapy. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins,
XX anti-bodies and (ant)agonists are useful in the diagnosis, treatment
XX and prevention of: (a) cancer, e.g. breast and ovarian cancer and
XX other cancers of the adrenal gland, bone, bone marrow, breast,
XX gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
XX (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
XX healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
XX and (f) infectious diseases such as viral, bacterial, fungal and
XX parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1,688-49 Length: 19125
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 21.66% Indels: 0
DB: 22 Gaps: 0
US-09-896-522-2 (1-277) x AAL36440 (1-19125)
QY 218 ValAlaIleAsnLeuIleValIleGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
Db 5997 GTGGCCATCACTGATCGTGCAGCACATCCAGACATTGTGATGTGTCATCTGCAA 6056
QY 238 TTPHIArGlyGlySerAsnGlyAArgSerTyrIlyAArgThrPheserGluProGlyAsp 257
Db 6057 TGGCACCAGAGAGAGTCCAAATGGGCGGAGCTACAAGCGAGACCTTTCTGAGCCAGGAGAC 6116

QY 258 HisProGlyMetLeuThrSerGlyLysAArgSerHisIleGluSerSerAArgProHis 277
Db 6117 CACCTGGATGCTGCTCTGGCAACCGGTACATTGGAGTCCAGACAGACCCAC 6176
RESULT 13
ID ABR59428 standard; cDNA; 19125 BP.
XX ABR59428;
AC
XX
XX 26-FEB-2003 (first entry)
XX
XX
XX cDNA encoding novel human musculoskeletal system antigen #1772.
DE
XX
XX
XX Gene, sg; musculoskeletal system antigen; cancer; metastasis;
XX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
XX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
XX post-operative tissue repair; limb regeneration; neuronal growth;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX AIDS-related complex; chondrocyte growth; bone regeneration;
XX periodontal regeneration; tissue transport; bone graft; skin aging;
XX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
XX cell growth; organ transplant; cell differentiation; body height;
XX weight; hair colour; eye colour; skin; percentage of adipose tissue;
XX pigmentation; cosmetic surgery; metabolism; biohythm; circadian rhythm;
XX depression; tendency for violence; pain; reproductive capability;
XX hormone level; endocrine level; appetite; lipid content; stress;
XX storage capability; fat content; lipid content; protein content;
XX carbohydrate content; vitamin content; cofactor content;
XX nutritional component.
XX
XX Homo sapiens.
OS
XX
XX US2002147140-A1.
XX
XX 10-OCT-2002.
XX
XX
XX 17-JAN-2001; 2001US-0764877.
XX
XX
XX 31-JAN-2000; 2000US-179065P.
XX 04-FEB-2000; 2000US-180628P.
XX 28-JUN-2000; 2000US-214886P.
XX 07-JUL-2000; 2000US-216647P.
XX 07-JUL-2000; 2000US-216680P.
XX 11-JUL-2000; 2000US-217487P.
XX 11-JUL-2000; 2000US-217496P.
XX 14-JUL-2000; 2000US-218290P.
XX 26-JUL-2000; 2000US-220963P.
XX 26-JUL-2000; 2000US-220964P.
XX 14-AUG-2000; 2000US-224518P.
XX 14-AUG-2000; 2000US-224519P.
XX 14-AUG-2000; 2000US-225267P.
XX 14-AUG-2000; 2000US-225268P.
XX 14-AUG-2000; 2000US-225270P.
XX 14-AUG-2000; 2000US-225447P.
XX 14-AUG-2000; 2000US-225757P.
XX 14-AUG-2000; 2000US-225758P.
XX 22-AUG-2000; 2000US-226868P.
XX 30-AUG-2000; 2000US-228924P.
XX 01-SEP-2000; 2000US-229287P.
XX 01-SEP-2000; 2000US-229343P.
XX 01-SEP-2000; 2000US-229344P.
XX 01-SEP-2000; 2000US-229345P.
XX 01-SEP-2000; 2000US-229509P.
XX 05-SEP-2000; 2000US-229513P.
XX 08-SEP-2000; 2000US-231413P.
XX 21-SEP-2000; 2000US-234223P.
XX 21-SEP-2000; 2000US-234274P.
XX 25-SEP-2000; 2000US-234997P.
XX 27-SEP-2000; 2000US-235834P.
XX 29-SEP-2000; 2000US-236327P.
XX 29-SEP-2000; 2000US-236367P.
XX 29-SEP-2000; 2000US-236368P.

PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 02-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239353P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.
 PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 PR 08-DEC-2000; 2000US-251869P.
 PA (ROSE/) ROSEN C A.
 PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI: 2003-128199/12.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 XX associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX
 PS Disclosure; SEQ ID NO 2805; 31pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid: stimulates re-vascularization of ischemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since FGF family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryonic stem cells, besides, hematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism, changes
 CC mammal's metal state or physical state by influencing biorythms,
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This sequence encodes a novel human
 CC musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.
 XX
 SQ Sequence 19125 BP; 4263 A; 5265 C; 5444 G; 4153 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,68e-49 Length: 19125
 Score: 60.00 Matches: 60
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 21.66% Indels: 0
 DB: 25 Gaps: 0
 US-09-896-522-2 (1-277) x ABX59428 (1-19125)
 Qy 218 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 237
 Db 5997 GTTGCCATCAACTGATCGTCGACGACATCCAGACATTCTGATGGTGCATCTGCAG 6056
 Qy 238 TrpHisArgGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluPProGlyAsp 257
 Db 6057 TGGCACCAGAGAGGTCCTCAATGGGCGAGCTAACAGCGACTTTTCTGAGCCAGGGGAC 6116
 Qy 258 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
 Db 6117 CACCTGGAGTGGTCTGACCTTGGCAACGGTCACTTTGAGTCCAGCAGACACCCAC 6176
 RESULT 14
 AA215454
 ID AA215454 standard; cDNA; 715 BP.
 XX
 AC AA215454;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 DE Human gene expression product cDNA sequence SEQ ID NO:2923.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN NO9938972-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99MO-US01619.
 XX
 PR 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 PI CkVenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;
 PI Jones WL, Kasam A, Kennedy GC, Kita D, Labat J;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX
 XX WPI: 1999-494092/41.
 DR
 XX Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 PT
 PS Claim 1; Page 1410-1411; 2479pp; English.
 XX
 CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AA21532 to AA21779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AA21532 to AA21779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and

CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein), and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX Sequence 715 BP; 179 A; 189 C; 178 G; 157 T; 12 other;

Alignment Scores:

Pred. No.:	1,44e-15	Length:	715
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.03%	Indels:	0
DB:	20	Gaps:	0

US-09-896-522-2 (1-277) x AA215454 (1-715)

QY 189 ThrPheValIysPProAlaPheGluGluPheCysIeuProThrIyblYrAlaaspVal 208
Db 470 ACgTtGtCAAGcTtGcCTTGAAGAATtCTGtTGCcAAAGAAgTATGCTGAATG 529

QY 209 ILeIlePProArgGly 213
Db 530 ATCATCCCTAGAGGT 544

RESULT 15

AAFI5785
ID AAF15785 standard; cDNA; 1310 BP.

XX AAF15785;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:220.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytosolic; cardiovascular; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
XX antiinfective; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease; ss.

OS Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENE SCI INC.
(ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI, 2000-587513/55.

DR P-PSDB; AAB56582.

XX Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

XX Claim 1, Page 752, 2338pp; English.

PS AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardiovascular, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 1310 BP; 338 A; 350 C; 352 G; 264 T; 6 other;

Alignment Scores:	2.54e-15	Length:	1310
Pred. No.:	25.00	Matches:	25
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	9.03%	Gaps:	0
DB:	21		

US-09-896-522-2 (1-277) x AAF15785 (1-1310)

QY 189 ThrPheValIysPProAlaPheGluGluPheCysIeuProThrIyblYrAlaaspVal 208
Db 789 ACgTtGtCAAGcTtGcCTTGAAGAATtCTGtTGCcAAAGAAgTATGCTGAATG 848

QY 209 ILeIlePProArgGly 213
Db 849 ATCATCCCTAGAGGT 863

Search completed: November 25, 2003, 10:20:43
Job time : 349 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 25, 2003, 09:29:29 ; Search time 3972 Seconds

(without alignments)
2852.962 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGDECPAPRAPH.....HPGLTRSGKSHLESSRPH 277

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Word size: 15
Total number of hits satisfying chosen parameters: 50

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09896522/runat.21112003.184459.5866/app_query.faeta.1.455
-DB=GenEmbl -QFMT=fastap -SUFIX=p2n.oli.rge -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PRO -NORM=ext -HEARST=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CGN.1.1.3508.@runat.21112003.184459.5866 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_da.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_hum.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pin.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	277	100.0	834	6	AX449219	AX449219 Sequence
2	277	100.0	834	9	AF254133	AF254133 Homo sapi
3	277	100.0	1022	9	AF237290	AF237290 Homo sapi
4	277	100.0	1624	6	AX449217	AX449217 Sequence
5	277	100.0	2160	6	BD157613	BD157613 Primer fo
6	277	100.0	2160	9	AK022317	AK022317 Homo sapi
7	191	69.0	753	6	BD146824	BD146824 Primer fo
8	189	68.2	1395	9	AF125106	AF125106 Homo sapi
9	169	61.0	2072	9	BC015547	BC015547 Homo sapi
10	165	59.6	2228	9	AK057848	AK057848 Homo sapi
11	112	40.4	734	6	AX540411	AX540411 Sequence
12	60	21.7	14792	9	AL358781	AL358781 Human DNA
13	45	16.2	1810	10	MUSURKI	L31783 Mus musculu
14	45	16.2	1959	10	BC025146	BC025146 Mus muscu
15	45	16.2	25493	2	AC078885	AC078885 Mus muscu
16	30	10.8	841	11	BV017303	BV017303 S212P6359
17	30	10.8	952	10	MUSURKIA	L31784 Mus musculu
18	30	10.8	192336	2	AC118474	AC118474 Mus muscu
19	30	10.8	211075	10	AC076974	AC076974 Mus muscu
20	25	9.0	314	10	AB030700	AB030700 Rattus no
21	25	9.0	336	9	BT006860	BT006860 Homo sapi
22	25	9.0	336	12	BT007548	BT007548 Synthetic
23	25	9.0	744	9	AB062451	AB062451 Homo sapi
24	25	9.0	1121	9	AF236637	AF236637 Homo sapi
25	25	9.0	1209	9	BC002906	BC002906 Homo sapi
26	25	9.0	1312	10	BC023789	BC023789 Mus muscu
27	25	9.0	1322	6	AX135546	AX135546 Sequence
28	25	9.0	1354	10	AF236636	AF236636 Mus muscu
29	25	9.0	1784	5	BC045968	BC045968 Danio rer
30	24	8.7	660	11	BV077668	BV077668 S212P6227
31	24	8.7	233210	2	AC097693	AC097693 Rattus no
32	24	8.7	241882	10	AL808027	AL808027 Mouse DNA
33	24	8.7	244105	2	AC098897	AC098897 Rattus no
34	24	8.7	254677	2	AC068494	AC068494 Rattus no
35	19	6.9	383	11	G41891	G41891 SHGC-35183
36	17	6.1	655	11	BV077340	BV077340 S212P6009
37	16	5.8	67374	2	AC118933	AC118933 Mus muscu
38	16	5.8	99866	2	AL358115	AL358115 Human DNA
39	16	5.8	120513	2	AC142244	AC142244 Mus muscu
40	16	5.8	148865	2	AL554171	AL554171 Danio rer
41	16	5.8	164246	2	AC016371	AC016371 Homo sapi
42	16	5.8	187700	2	AC131061	AC131061 Mus muscu
43	16	5.8	233137	2	AC127758	AC127758 Rattus no
44	16	5.8	236660	2	AC097934	AC097934 Rattus no
45	15	5.4	831	9	D78335	D78335 Human mRNA

RESULT 1

ALIGNMENTS

AX449219
LOCUS AX449219 834 bp DNA linear PAT 03-JUN-2002
DEFINITION Sequence 3 from Patent WO0202761.
ACCESSION AX449219
VERSION AX449219.1 GI:21697996
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gluckmann, M.A.
TITLE 57658: a human uridine kinase and uses thereof
JOURNAL Patent: WO 0202761-A 3 10-JAN-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1. 834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 194 a 218 c 265 g 157 t
ORIGIN
Alignment Scores: 2.6e-298 Length: 834
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x AX449219 (1-834)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40
Db 61 CAGGGCCCTTCGATGAGGGTGGAGCGCGGCACTCCAGCGGGAAGTCGACGTTGT 120
QY 41 GluValIleMetGluLeuGluGlyValAsnGluValGluGlnAspGlnArgValVal 60
Db 121 GAGAGATCATGAGTCTGCTGGGACAGACGAGGTGAAACGCGGACGAGGAGTGTCTC 180
QY 61 IleLeuSerGlnAspArgPheGlyValValLeuThrAlaGluGlnValAlaValAlaLeu 80
Db 181 ATCTGAGCCAGGACAGGATCTCAAGGTCTCTGACGCGACAGGACAGGAGGCTTG 240
QY 81 LysGlyGlnThrAsnProPheAspHisProAspAlaPheAspAsnProLeuMetHisArgThr 100
Db 241 AAAGAGACAGTCAATTTTGGACATCCAGATGCCCTTTGATTAATGATTGATGACAGACT 300
QY 101 LeuValAsnIleValGluGlyValThrValGluValProThrTyrAspPheValThrHis 120
Db 301 CTGAGAGACATCTGCGAGGGGCAAAACGTTGAGAGTCCGACTATGATTTTGTACACAC 360
QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
Db 361 TCAGGTTTACAGAGACACGAGGTGTCTACCTCGGACGTTGTTGTTGAGGAGATC 420
QY 141 LeuValPheTyrSerGlnGluIleArgAspMetHehIleuValGluPheValAspThr 160
Db 421 TTGGTGTTCACACCCAGGAGATCCGGGACATGTTTCCACCTGCCCTCTCTGTGAGACC 480
QY 161 AspSerAspValAlaGluSerArgArgValLeuArgAspValAlaGAGGlyArgAspLeu 180
Db 481 GACTCCGACGTCAGAGCTGTCTCGAAGAGTTTCCGGGAGCTGCCCGAGGGAGGAGCTG 540
QY 181 GluGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu 200
Db 541 GAGCAGATTTCAGACGACGTACACCACTTGTGTAAGCGGCGCTTCGAGGAGTTTGTGCTG 600

QY 201 ProThrIleuValSerThrAlaAspValIleIleProArgGlyValAlaAspAsnMetValAlaIle 220
Db 601 CCAGCAAGAGATATCCAGATGATCATCCACGAGAGGTGACATATGTTGGCATC 660
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleThrHisArg 240
Db 661 AACCTGATGTCGAGACATCCAGACATTTCTGAATGGTGACATCTGCMAATGACACCA 720
QY 241 GlyIleSerAsnGlyArgSerTyrIleValArgThrPheSerGluProGluAspHisProGly 260
Db 721 GGAGGCTCCATGCGGAGCTCAAGCGGACCTTTCTGAGCCAGGAGGACCACTCGG 780
QY 261 MetLeuThrSerGlyValArgSerHisIleuGluSerSerArgProHis 277
Db 781 ATGCTACCTCTGGCAAGCGGTCACTTGGATCCAGACGAGACCCAC 831
RESULT 2
LOCUS AF254133 834 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens uridine kinase mRNA, complete cds.
ACCESSION AF254133
VERSION AF254133.1 GI:13924749
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Ho, Y.S. and Johnson, R.K.
TITLE 1 (bases 1 to 834)
JOURNAL Human uridine kinase from prostate cancer cell line (LNCap)
REFERENCE 2 (bases 1 to 834)
AUTHORS Ho, Y.S. and Johnson, R.K.
TITLE Direct Submision
JOURNAL Submitted (10-APR-2000) Oncology Research, SmithKline Beecham, 709
Swedeland Road, King of Prussia, PA 19406, USA
FEATURES
source 1. 834
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 834
/note="protein kinase"
/codon_start=1
/product="uridine kinase"
/protein_id="AAK49122.1"
/db_xref="GI:13924750"
/translation="MASAGDEDCESPAPEADRPHPQPLIGVSGTASGKSTVCEKIM
ELIGONEVROBOROKVITLSODRKYKVLAEOKAKLKGQYNDHPDPAFNDIMHRTLK
NIVEGTIVTPYDPTVTHSRLEPTTVYFADVVLPGILIVFSQELRDMFHLRIPTDT
DSVRLSRKVLVDYRKRGRDLBOILTYTTTFVKPAPEFELPTKKYADVILIPRGVDMV
AINLVQHIQDILNGDICRKHGSGNSGRSYKRTFSEPHGPGLTSGKRSHTLESSRP
H"
BASE COUNT 194 a 219 c 265 g 156 t
ORIGIN
Alignment Scores: 2.6e-298 Length: 834
Score: 277.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-896-522-2 (1-277) x AF254133 (1-834)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db 1 ATGGCTTCGGGGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGACCGTCCGAC 60
QY 21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys 40
Db 61 CAGGGCCCTTCGATGAGGGTGGAGCGCGGCACTCCAGCGGGAAGTCGACGTTGT 120

Qy	41	Glulys11IemctGluLeuLeuGlyGlnasnGluVal1GluGlnInargGlnInargLysValVal1	60
Db	121	GAGAGAGATCATGAGAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACAGGAAAGGTGCTC	180
Qy	61	11LeuSerGln1nAspArgPheTyrCysVal1LeuThr1aGluGln1aLysVal1aLeu	80
Db	181	ATCTGAGCCACGAGACAGGTTTCTACAGGTTCTTGACGCGACAGACAGAGGCCAGGCTTG	240
Qy	81	LysGlyGlnTyrAsnPheAspHisProAsp1aPheAspAsnAspLeuMetHisArgThr	100
Db	241	AAAGGACAGTCAATTGTCACCATTCAGATGCTTGAACAATGATTTGATGCACAGACT	300
Qy	101	LeuLysAsn11LeValGluGlyLysThrPalGluVal1ProThrTyrAspPheValThHis	120
Db	301	CTGAAGAACATCGTGGAGGCGCAAAACGGTGGAGGGCGCCAGCTTAATGATTTGTGCACAC	360
Qy	121	SerArgLeuProGluThrThrValValTyrPro1aAspVal1ValLeuPheGluGly1Le	140
Db	361	TCAAGGTTACAGAGAACCAACGGTGGTCTACCTCGGAGACGTGTTCTGTGGAGGCATC	420
Qy	141	LeuValPheTyrSerGlnGlu11LeArgAspMetPheHisLeuArgLeuPheValAspThr	160
Db	421	TTGGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCACTCGCGCCTCTTGTTGAGACAC	480
Qy	161	AspSerAspValArgLeuSerArgArgVal1LeuArgAspValArgArgGlyArgAspLeu	180
Db	481	GACTCCGACGTACAGGCTCTCTGAAAGATTTCTCCGGGACGTGCGCCGAGGAGGGACCTG	540
Qy	181	GluGln11LeuThrGlnTyrThrThrPheValLysPro1aPheGluGluPheCysLeu	200
Db	541	GAGCAGATTCGACGCACTACACCACTTCCTGGAGACCGGCTTCGAGAGATTCGCTG	600
Qy	201	ProThrLysLysTyrAlaAspVal11Le1ProArgGlyValAspAsnMetValAla1Le	220
Db	601	CCGACAAAGAAAGTATGCCGATGTGTATCCACACAGAGAGTGAGCAATATGTTGCCATC	660
Qy	221	AsnLeu11LeValGlnHis11LeGlnAsp11LeuAsnGlyAsp11LeCysLeuThrHisArg	240
Db	661	AACCTGATCGACACACATCCACAGACATTCCTGAATGGTGAATCTGCAAAATGGACCGA	720
Qy	241	GlyLysSerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly	260
Db	721	GGAAGGTCATATGGCGGAGGCTACAGAGGACCTTTTCTGAGCCAGGGGACCACTCTGG	780
Qy	261	MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis	277
Db	781	ATGCTGACCTCTGGCAAAACGGTCACTTTGGAGTCCACAGACAGACCCAC	831
RESULT 3			
AF237290			
LOCUS	AF237290	1022 bp	mRNA linear PRI 02-AUG-2001
DEFINITION	Homo sapiens uridine-cytidine kinase 1 (UCK1) mRNA, complete cds.		
ACCESSION	AF237290		
VERSION	AF237290.1	GI:13506764	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1022)		
AUTHORS	Van Rompay,A.R., Norda,A., Linden,K., Johansson,M. and Karlsson,A.		
TITLE	Phosphorylation of uridine and cytidine nucleoside analogs by two		
JOURNAL	human uridine-cytidine kinases		
MEDLINE	Wol. Pharmacol. 59 (5), 1181-1186 (2001)		
PUBMED	21203813		
REFERENCE	11306702		
AUTHORS	2 (bases 1 to 1022)		
	Van Rompay,A.R., Linden,K., Norda,A., Zhu,C., Zheng,X.,		
	Johansson,M. and Karlsson,A.		
TITLE	Human uridine-cytidine kinase 1 and 2 : rate limiting enzymes		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 1022)		

AUTHORS	Van Rompay, A. R., Linden, K., Norda, A., Zhu, C., Zheng, X., Johansson, M., and Karlsson, A.
TITLE	Direct Submission
JOURNAL	Submitted (19-FEB-2000) IMPI, Clinical Virology, Huddinge University Hospital, Stockholm 14186, Sweden
FEATURES	Location/Qualifiers
SOURCE	1. .1022
gene	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"
CDS	1. .1022 /gene="UCK1" 5. .838 /feature="UCK1" /function="phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate" /codon_start=1 /product="uridine-cytidine kinase 1" /protein_id="AAK28324.1" /db_xref="GI:13506765" /translation="MASAGGEDCESPAEDRPHQRFPLIGVSGTASGKSTCEKIMELGQNEVEORQRKVILISODRPFYKVLTAQKAKALGOVNFDPDIFNDIMRTLKIMVEGKTEVEPVDVDFRHSRLPETTVYPADVVFEEGILVFSOEDIMFLRLEPDMDSVDRLSRRLVDRRGRDLBOLITQYTTVPKPAFEESCLPTKKYADVIIIRGVDMWAINLIYHIQDILANGDICKMRGSGSNRSYKRTSBEPDHPGMLTSGKSHLESSRPH"
BASE COUNT	227 a 282 c 324 g 189 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.16e-298 Length: 1022
Score:	277.00 Matches: 277
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	9 Gaps: 0
US-09-896-522-2 (1-277) x AF237290 (1-1022)	
QY	1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db	5 ATGGCTTGGGGGAGGCGAAGACTGGAGAGCCCCGGCGGAGCCGACGATCCGCAC 64
QY	21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
Db	65 CAGCGGCCCTTCCGATAGGGGTAGGGCGGGCGGCACTCCAGCGGGAATGCACCGTGT 124
QY	41 GluIleValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
Db	125 GAGAAAGATCATGAGGTGCTGGGACAAAGACAGGTGGAACAGCGGCGAGCGGAAGTGGTC 184
QY	61 IleIleuSerGlnAspArgPheThrLysValIleuThrLysGluGlnIleuValAlaValAlaIleu 80
Db	185 ATCTGTGACCAAGGACAGAGTTCCTTACAAAGTCTCTGACGCGACAGCAAGAGGCCAAAGCCCTTG 244
QY	81 LysGlyGlnThrLysPheAspPheAspProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db	245 AAAAGACAGTACATTTTGAACATCCAGATGCCCTTTGATTAATGATTTGGATGCAACAGACT 304
QY	101 LeuLysAsnIleValGluGlyLysThrValGluValProThrLysAspPheValThrHis 120
Db	305 CTGAAGAAACATCGTGAAGGCAAAACGGTGAAGGTGCCGACCTATGATTTTGTGACACAC 364
QY	121 SerArgLeuProGluThrThrValValLysThrProAlaAspValValLeuPheGluGlyTyr 140
Db	365 TCAAGGTTTACACAGACCAACGGGTGTACCTGCGGACGCGTTCGTGTTGAAGGCAATC 424
QY	141 LeuValPheThrLysSerGlnGluIleArgAspMetPheHisLeuArgLysPheValAspThr 160
Db	425 TTGGTGTTCACACCCGAGAGATCCGGGACATGTTCCACTGCGCCCTTCCTGTTGACACAC 484
QY	161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180

Db	485	GACCTCCGACGTCAAGCTGTCTCGAAGAAGTTCTCCGGAGACTGGCCGAGAGGAAGGCACCTGT	544
Oy	181	GlucInIIeleuthrGlnTyTrThrPheValIAspProAlaPheGIugluPhCySleu	200
Db	545	GAGCAGATTCTGACGCGTAGTACACCACCTTCGTGAAGCGGCCCTTCGAGGAGTTCTGCCGTG	604
Oy	201	ProthrIysIserTyTrAlaaspValIleIleProArgGlyValAspMetValAlaIle	220
Db	605	CCGACAAAGAAAGTATGCCGATGTGATATCCCCGAGAGAGTGACAATATGTGTGCATC	664
Oy	221	AsnIeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCySylTrpHisArg	240
Db	665	AACCTGATCGCGACGACATCCAGGACATTTCTGAATGTGACATCTGCCAAATGGCACCCA	724
Oy	241	GIyGIySerAsnGIyArGserTyTrLysArgTrh-PheSerGIuDPGCIyAspHisProGly	260
Db	725	GGAGGGTTCATATGGCGGAGCTACAAACGACCTTTTCTGACGACGAGGACACCTCTGG	784
Oy	261	MelIeuThrSerGIyLyAsrGserHisIeUGIuSerSerSerArgProHis	277
Db	785	ATGCTGACCTTCGGCAAACGGTCACTTTGGAGTCCAGACGAGACCCAC	835
RESULT_4			
AX449217	AX449217	1624 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 1 from Patent WO0202761.		
ACCESSION	AX449217		
VERSION	AX449217.1 GI:21697994		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Gluckemann,M.A.		
TITLE	57658, a human uridine kinase and uses thereof		
JOURNAL	Patent: WO 0202761-A 1 10-JAN-2002;		
FEATURES	Millennium Pharmaceuticals, Inc. (US) Location/Qualifiers 1..1624 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
CDS	94..927 /note="unnamed protein product" /codon_start=1 /protein_id="CADJ7771.1" /db_xref="GI:21697995" /translation="MASAGEDECSPEAPEDRPHORPELLIGVSGTAGSKSTVCCKIM ELAQNEVEQQRKVLLSQDRFYKVLTAQKKAKALKGYNPDHPADFNDAHTLK NIVGKTVEVTYDFTHSRLEPTFVADVLFEGLIFVSOEIRDFHLFLFDI DSVDRLRRVLRDVRGRDELLOITTYTFVKAFEPFCIPTKKADVILPRGVDMV AINLIYHIQDIINGDICMKWHRGSSNRSLKRTFSRPGHPLMLISGKRSHLESSRP H"		
BASE COUNT	354 a 427 c 506 g 337 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4,92e-298	Length:	1624
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6 Gaps:		0
US-09-896-522-2 (1-277) x AX449217 (1-1624)			
Oy	1	MetaIaseralagIyGIyUASPcGylUserProAlaprogIuaIlaaspArgProHis	20
Db	94	ATGGCTTCGGGGAGGAGGAAGACTGGAGAGCCCGCGCGAGGCCAGCCTCCGCAC	153
Oy	21	GlnArgProPheIeuIleGlyValSerGIyGIyThraIaSerGIyIysSerThraValCys	40

```

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES
    source          Location/Qualifiers
                    1. 2160
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT      457 a 591 c 671 g 441 t
ORIGIN
Alignment Scores:
Pred. No.:      6,48e-298      Length:      2160
Score:          277.00         Matches:      277
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%         Indels:      0
DB:            6              Gaps:          0
US-09-896-522-2 (1-277) x BD157613 (1-2160)
QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
DB 95 ATGGAGCTTCGGCGGAGGCGGAAGACTCGGAGAGCCCGCGCGGAGCGCCGTCGCAC 154
QY 21 GlnAyrProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
DB 155 CAGGGGCCCTTCCTGATAGGGGTAGGGGGCGGCGCTCCAGCGGGAAGTGCACCGTGTGT 214
QY 41 GlnLysIleMetGluLeuGluGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
DB 215 GAGAGATCATGAGGTGCTGGGACAGAAAGAGGTGAAACAGCGGCGAGCGGAGAGTGTCTC 274
QY 61 IleLeuSerGlnAspArgPheTyrrLysValLeuThrAlaGluGlnValAlaLysAlaLeu 80
DB 275 ATCTGTGCGCAGGACGAGTCTCAACAGTCTCTGACGCGCAGGACAGAGCGCCCTTG 334
QY 81 LysGlyValIntyrAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
DB 335 AAGAGACAGTACATTTTGAACATCCAGATGCCCTTATATATGATTTGATGCACAGGACT 394
QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrrAspPheValThrHis 120
DB 395 CTGAAGAACATCTCGTGGAGGGCCAAACCGTGGAGGTGCCGACCTATGATTTTGTGACACAC 454
QY 121 SerArgLeuProGluThrThrValValTyrrProAlaAspValValLeuPheGluGlyIle 140
DB 455 TCAAGGTATCACAGAGCACCGTGTCTACCTCGGACGCTGTCTGTTTGAAGGCGATC 514
QY 141 LeuValPheTyrrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
DB 515 TTGGTGTCTTCAACGCGAGGATCCGGAGCATGTTCACCTGCGCTCTTCTGTGAGACACC 574
QY 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
DB 575 GACTCCGACATCGAGCGCTGTCTCGAAGAGTCTCCGGGACGTCGCCGACGAGGAGGACCTG 634
QY 181 GlnGlnIleLeuThrGlnTyrrThrPheValLysProAlaPheGluGluPheCysLeu 200
DB 635 GAGCAGATTCCTGACGCGAGTACACACACTTGTGTGAAGCGGCGCTCGAGAGTCTGCTG 654
QY 201 ProThrLysLysTyrrAlaAspValIleIleProArgLysValAspAsnMetValAlaIle 220
DB 695 CCGACAAAGAGATGTGCGATGTGATCATCCACGAGAGGTGACCAATATGTTGCCATC 754
QY 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240

```

```

DB 755 AACCTGATCGTGCAGACATCCAGACATTCGAATGTGATCATCTGCAAAATGACCGCA 814
QY 241 GlnGlySerAsnGlyArgSerTyrrLysArgThrPheSerGluProGluAspHisProGly 260
DB 815 GAGAGGTTCATATGGCGGAGCTTCAAGCCGACCTTTTCTGAGCCAGGAGCACCCCTGGG 874
QY 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
DB 875 ATCTGACCTCTTCGCAACGCTACATTTGAGTCTCAGACGACGACCCAC 925
RESULT 6
AK022317      2160 bp      mRNA      linear      PRI 01-AUG-2002
LOCUS        Homo sapiens cDNA FLJ12255 fis clone MAMMAL001476, highly similar
DEFINITION   to URIDINE KINASE (EC 2.7.1.48).
ACCESSION   AK022317
VERSION     AK022317.1 GI:10433687
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    1
AUTHORS     Isogai,T., Ota,T., Hayaeshi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
            Magatsuna,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
            Ono,Y., Takiguchi,S., Watanabe,S., Kinura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,U., Wakamatsu,A.,
            Nakamura,Y., Nagahara,K., Masubo,Y., Ninomiya,K. and Iwayanagi,T.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 2160)
            Isogai,T. and Otsuki,T.
            Direct Submission
            Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
            Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            International Trade and Industry of Japan; cDNA full insert
            sequencing; Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection;
            Helix Research Institute (supported by Japan Key Technology Center
            etc.) and Department of Virology, Institute of Medical Science,
            University of Tokyo.
            Location/Qualifiers
            1. 2160
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="MAMMAL001476"
                /issue_type="Mammary gland"
                /clone_lib="MAMMAL"
                /note="Cloning vector: pME186FL3"
                95..928
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="BAB14010.1"
                /db_xref="GI:10433688"
                /translation="MASAGDECPAPAPDRPHORPFLIGVSGGTASGKSTVCEKIM
                ELLGNEVROROKRVILISODRFYVLTAEOKAKALKGOYNDHPDAPNDMLHRLTK
                NIVEGTVPEVPTDYTHSRLEPTTVYVPADYVLEBGLVFPISQELRDPFHLRLPDT
                DSDVLRSLVDVRRGLDQILTQYTFVVPAREFELPTPKTADAVILPFGVDNMV
                AINLVQHLQDILINDICKWHRGSNGRSYKRTFSEPDHPMLTSGKSHLESSRP
                H"
BASE COUNT      457 a 591 c 671 g 441 t
ORIGIN
Alignment Scores:
Pred. No.:      6,48e-298      Length:      2160
Score:          277.00         Matches:      277
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0

```

Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-896-522-2 (1-277) x AK022317 (1-2160)			
QY	1 MetAlaSerAlaGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis	20	
DB	95 ATGGCTTCGGCCGGAGGCGAAGCTGCGAGAGCCCGCCGAGGCGCAGCGTCCGCAC	154	
QY	21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyValSerThrValCys	40	
DB	155 CAGCGGCCCTTCTGATTAGGGGTGACCGCGGCACTGCCAGGGGAAGTCGACCGGTGT	214	
QY	41 GlyValIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal	60	
DB	215 GAGAAAGATCATGAGAGTTCCTGGGGAAGAAACGAGGTGAAACAGCGGACGGAAGGTGCT	274	
QY	61 IleLeuSerGlnAspArgPheThrValIleThrAlaGluGlnValAlaValAlaIleu	80	
DB	275 ATCCGAGGCCAGAGACGAGTTCCTCAAGGCTCCGACCGGACAGACAGAGGCCAAGCCCTTG	334	
QY	81 LysGlyGlnIleThrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100	
DB	335 AAGAGCAGTACAAATTTTGACATCCAGATGCCCTTGAAAGATTTGATGACAGGACT	394	
QY	101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyraAspPheValThrHis	120	
DB	395 CTGAAGAATCATGCTGGAGGGCAGAAACGGTGGAGAGTCCGACCTTAGATTTTTGACACAC	454	
QY	121 SerArgLeuProGluThrThrValValTyrProAlaAspValIleuPheGluGlyIle	140	
DB	455 TCAGAGTTTACAGAGACCAAGGTGGTCTACCTGCGAGAGTGTCTGTTTGAAGGCGATC	514	
QY	141 LeuValIlePheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160	
DB	515 TTGGTTTCTACAGCCAGAGAAATCCGGAGCATGTTCCACTCGGCTCTTCGTGACACC	574	
QY	161 AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyTyrAspLeu	180	
DB	575 GACTCGACAGTCAGAGGTGCTCGAAGAGTTCCTCCGGAGGTGCTCCCGAGGAGGAGACTG	634	
QY	181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGluPheCysLeu	200	
DB	635 GAGCAGATTCTGACCCAGACACCACTTCGTGAAGCCGGCTTGAGGAGTTCGTGCTG	694	
QY	201 ProThrLysLysTyrAlaAspValIlelleProArgGlyValAspAsnMetValAlaIle	220	
DB	695 CCGAAGAAAGAAATATGATCCGATGTGATCATCCACAGAGAGTGGAGAAATATGTTTCCATC	754	
QY	221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIleThrPheHisArg	240	
DB	755 AACCTGATCTGACAGACATCCAGACATTCGTAATGTGACATCTCGAAATGGACAGGA	814	
QY	241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly	260	
DB	815 GGAGGGTCCAAATGGGCGAGGTACAGAGGCACTTTTTCAGCCAGGGGGAACCACTTGG	874	
QY	261 MetLeuThrSerGlyLysArgSerHisIleuGluSerSerSerArgProHis	277	
DB	875 ATGTGACCTCTGGCAAAACGGTCAACATTTGGATGTCAGAGACAGACCCAC	925	
RESULT 7			
LOCUS	BD146824	753 bp	DNA linear
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		
ACCESSION	BD146824		
VERSION	BD146824.1	GI:27852582	
KEYWORDS	JP 2002191363-A/1667.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 753)		

AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T. Primer for synthesizing full-length cDNA and use thereof Patient: JP 2002191363-A 1667 09-JUL-2002;									
TITLE	HELIIX RESEARCH INSTITUTE									
JOURNAL	OS Homo sapiens (human) PN JP 2002191363-A/1667 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU PI KEIICHI NAGAI,TETSUJI OTSUKI PC									
COMMENT	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key									
FEATURES	Location/Qualifiers 1..753 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"									
source	FT source 1..753 FT /organism="Homo sapiens (human)"									
BASE COUNT	158 a 196 c 261 g 135 t 3 others									
ORIGIN										
Alignment Scores:	Pred. No.: 1.39e-202 Length: 753 Score: 191.00 Matches: 191 Percent Similarity: 100.00% Conservative: 0 Best Local Similarity: 100.00% Mismatches: 0 Query Match: 68.95% Indels: 0 DB: Gaps: 0									
US-09-896-522-2 (1-277) x BD146824 (1-753)										
QY	1 Meta1aSer1aAG1gY1gY1uAaPc6e1uSerPro1aAPROG1uA1aAaPArgPro1h1s 20									
Db	95 ATGGCTTGGGGGGGGGAGGAGCACTGCAGAGCCCCGGGGCGGAGGCGCAGCTCCGAC 154									
QY	21 G1nArGPpRoPeLeu11eG1Vva1SeRg1gY1rTh1a1SeRg1LyVeSeRTh1Va1CyS 40									
Db	155 CAGCGGCCCTTCTCGATAGGGGTGAAGCCGGCGGCACTGCCAGCGGAAATCGACCTGTGT 214									
QY	41 G1uYr1s1eMeTg1uLeu1eG1yG1naSnG1uVa1G1uG1nArG1nArG1nArG1yVa1 60									
Db	215 GAGAAGATCA1CGAGTTCTCGGACAGAACAGAGTGGAACAGCGGCACGGAGAGTGTCTC 274									
QY	61 11e1eUsErG1nAaPArgPheTy1LyVa1LeuTh1a1AG1uG1nLyVa1a1eU 80									
Db	275 ATCTGAGCCAGGACAGGTTCTTACAGGCTCTTGACGACGACGAGCGCAAGGCCCTTG 334									
QY	81 LysG1gY1nTyraNpHeaSPH1sProaSP1aPheAaPaNaAPLeUeCh1aRgThr 100									
Db	335 AAAGGACG1CAAA1TTTGACCACTCGATGCTTTGATTAATGATTGATGACACAGACT 394									
QY	101 LeuY1aSn11eVa1G1uG1Ly1eThx1a1G1uVa1ProTh1TyAaSPpHeVa1Th1h1s 120									
Db	395 CTGAAGAACATCGTGGAGGCGAAAACGTTGAGGTCCGCACTATGATTTGTGACACAC 454									
QY	121 SerArg1eUpProG1uThTh1Va1Va1TyRPro1aAaPVal1a1eUpHeG1uG1Ly1Le 140									
Db	455 TCAAGGTTACACAGAGACCACGGTGTCACTCGCGGACGTGTCTGTGGAGGCGATC 514									
QY	141 LeuVa1PheTy1rSeRg1nG1u11eArGAspHeTh1s1eUaNg1eUpHeVa1AaSPHr 160									
Db	515 TTGGTGTTTCAACCGCAGAGATCCGGGACATGTTCCACTCGCGCTTTGTGTGACACC 574									
QY	161 AspSerAaPVal1Arg1eUsErArg1aGVal1LeuArGAsPVal1ArgArgG1YArgSPLeu 180									

Db 575 GACTCCGACGTCAGGTCTCGAAGATTCTCCGGACGCGCGAGGAGGACCTG 634

Qy 181 GluGlnIleuThrGlnTyrThrThrpheVal 191

Db 635 GAGCAGATTCTGACGACGACACCACTTGCTG 667

RESULT 8

AF125106 1395 bp mRNA linear PRI 07-FEB-2002

LOCUS AF125106 Homo sapiens uridine kinase mRNA, complete cds.

DEFINITION AF125106 Homo sapiens uridine kinase mRNA, complete cds.

ACCESSION AF125106

VERSION AF125106.1 GI:18568108

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 2 Xln,Y.R., Yu,L. and Zhao,S.Y. Cloning of a new human cDNA similar to Mus musculus uridine kinase mRNA

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1395)

AUTHORS Ding,J.B., Yu,L. and Zhao,S.Y.

TITLE Direct Submission

JOURNAL Submitted (02-FEB-1999) Lab of Human Gene Research, Institute of Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433, People's Republic of China

FEATURES

source

1. 1395

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

8. 838

/codon_start=1

/product="uridine kinase"

/protein_id="AAL75943.1"

/db_xref="GI:18568109"

/translation="MASAGEGARAARARARPPHRLPGVSGTASGKSTVEKIKE LLDQNEVEQRHGVILISQRFYKVLAEKAKALKQYVDPDHPDADNDLMHRTLKN IVEGKTVEPTVDFVTHSRLEPTVVPAADVAFEGILVPSQIRDMFHLRFVDID SDVLSRRVLRVDVRRGRLEQLITQYTFKPAFEERCLPTKXADVIIPRGVNMVA INLIHQIIDLNDICKMRHSGNSGRVYKRTSEPDHDMGLTSGKSHLESSRHH "

CDS

BASE COUNT 325 a 352 c 411 g 307 t

ORIGIN

Alignment Scores:

Pred. No.: 4.24e-200 Length: 1395

Score: 189.00 Matches: 258

Percent Similarity: 98.47% Conservative: 0

Best Local Similarity: 98.47% Mismatches: 2

Query Match: 68.23% Indels: 4

Gaps: 0

US-09-896-522-2 (1-277) x AF125106 (1-1395)

Qy 18 ArgProHisGlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSer 37

Db 56 CGTCCGACACGAGCGTCCCTTCCTATAGGGGTGACGCGGCACTGCCAGCGGAACTCG 115

Qy 38 ThrValGlyGlyValIleMetGluLeuLeuGlyGlnIleGluValGluGlnArg 57

Db 116 ACCGTGTGTGAAGATCATGAGTTCCTGGGACAGACAGAGGTGGAACAGCGGCA-CGG 174

Qy 58 -LysValValIleLeuSerGlnAspArgPheTyrLysValIleuThrAlaGluGlnLysAl 77

Db 175 TAGAGTGATATCTCTGACGACGACAGAGTTCACAGTCTTGAACGCGACAGAGAGGC 234

Qy 77 AlaValAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 97

Db 235 CAAGGCGCTTGAAGACAGTACATTTTGACATCCAGATGCCCTTGATATGATTTGAT 294

Qy 97 cHisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspH 117

Db 295 GCACAGACTCTGAAGACATCTGTGAGGGCAAAAGGTGAGGTGCCACCTATATTT 354

Qy 117 eValIrrHIsSerArgLeuProGluThrValValIlyrProAlaAspValIleuPh 137

Db 355 TGTACACACTCAAGGTTCACAGAGACCAAGGTGTCTACCTCGCGGACGTGTTCTGTT 414

Qy 137 eGluGlyIleLeuValPheTyrSerGlnGluIleAlaGAspMetPheHisLeuArgLeuPh 157

Db 415 TGAAGGCATCTTGTTGTCTTACAGCCAGAGATCCCGGACATGTTCCACTGGGCTCTT 474

Qy 157 eValAspThrAspSerAspValAlaArgLeuSerArgValIleuArgAspValAlaArgyl 177

Db 475 CGTGGACACCGACTCCGACGTCAGGCTGTCTCGAAGATTTCTCGGAGCGTGGCGCGG 534

Qy 177 yArgAspLeuGluGlnIleuThrGlnTyrThrThrpheValIysProAlaPheGluGl 197

Db 535 GAGGAGACTGGAGCAGATTCTGACGAGTACACCACTTCGTGAACCGGCTTCGAGGA 594

Qy 197 upheCyLeuProThrLysLysTyrAlaAspValIlelleProArgGlyValAspAsn 217

Db 595 GTTCTGCTGCCGACAAAGAGATGCCATGTGATCATCCACGAGAGTGACAAATAT 654

Qy 217 tValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysly 237

Db 655 GGTTCGACATCACTTATGTCGACATCCGACATTTCTGATAGTGACATCTGCAA 714

Qy 237 sTrpHisArgGlyGlySerAsnGlyArgSer-TyrLysArgThrPheSerGluProGly 257

Db 715 ATGGACCCGAGGAGGGTCCATGAGGGGAC-GRACAAGGGAGACTTTTTCGACCAAGGGG 773

Qy 257 sPHisProGlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProH 277

Db 774 ACCACCTGGAGATGACCTCTGCGCAACGATCATTTGAGTCCAGACGACACCCC 833

Qy 277 ts 277

Db 834 AC 835

RESULT 9

BC015547

LOCUS BC015547 Homo sapiens, similar to uridine-cytidine kinase 1, clone MGC:9668

DEFINITION IMAGE:3845821, mRNA, complete cds.

ACCESSION BC015547

VERSION BC015547.1 GI:15930229

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 2072)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.sbgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxli.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/JMLN, at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: f Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 13899252.

FEATURES

SOURCE

Location/Qualifiers
 1..2072

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9668 IMAGE:3845821"
 /tissue_type="Colon, adenocarcinoma"
 /clone_lib="NH MGC_65"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 87..692
 /product="similar to uridine-cytidine kinase 1"
 /protein_id="AAH15547.1"
 /db_xref="GI:15930230"
 /translation="MASAGEDCESPAPADRPQRPFLIGSGTASGKSTCEKIM
 ELIGONEVOROKRVILISODRFYKVLAEQKAKLQVAFDNDLMDHRTLK
 NIVEKTYEPTVDVTHSRLLPETHVYPAPVULPEGLIYPSGEIRDMFLRLFVLT
 DSDVRLSRKDEKRCDFPSTRSGYGCCHQPDRAHPHSEW"

CDS

BASE COUNT 455 a 559 c 637 g 421 t
 ORIGIN

Alignment Scores:

Pred. No.: 1..166-177 Length: 2072
 Score: 169.00 Matches: 169
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.01% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x BC015547 (1-2072)

QY 1 MetAlaSerAlaGlyGlyAlaSPcGyluSerProAlaProGluAlaAspArgProHis 20
 Db 87 ATGGCTTCGGCGGAGCGGAAGACTGCGAGAGCCCGCGGAGCGGACGTCGCCGAC 146
 QY 21 GlnArgProPheLeuIleGlyValSerGlyThrAlaSerGlyLysSerThrValCys 40
 Db 147 CAGCGGCCCTTCGATAGGGGTGAGCGGCGACATGCGACGGGAATCGACCGTGTGT 206
 QY 41 GluLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 Db 207 GAGAAAGATCATGAGTCTCGGACAGAACAGAGTGGAAACGCGGACGGGAAGTGTGTC 266
 QY 61 IleLeuSerGlnAspArgPheTyrIleValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 267 ATCTGAGCCAGGACAGGTTCTACAAAGTCTCGACGGCAGAGCAAGAGCCCTTG 326
 QY 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 327 AAAGACAGTCAATTGGACCATCCAGATGCTTTGATTAATGATTGATTCACAGAGACT 386
 QY 101 LeuLysAsnIleValGluGlyLeuThrValGluValProThrTyrAspPheValThrHis 120
 Db 387 CTGAAGAAATCATCGTGGAGGGCAAAACGTTGAGGTGCCGACCTTATGATTTTGTACACAC 446
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 447 TCAAGGTATACAGAGACCAAGGTGTCTACCTCGGACGTCGTCTGTTTGAAGGGATC 506
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuValGluLeuPheValAspThr 160
 Db 507 TTGTGTTCTACAGCCAGAGATCCGGGACATGTTTCCACTGCGCTCTTGTGTGACACC 566
 QY 161 AspSerAspValArgLeuSerArgArg 169
 Db 567 GACTCCAGCTCAGGCTTCTCGAAGA 593

RESULT 10

AK057848

LOCUS AK057848 2228 bp mRNA linear PRI 31-OCT-2001
 DEFINITION Homo sapiens cDNA FLJ25119 f1s, clone CBR05878, highly similar to

URIDINE KINASE (EC 2.7.1.48).

ACCESSION

VERSION

AK057848.1 GI:16553809

KEYWORDS

oligo capping; f1s (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isegaki, T. and Sugano, S.

TITLE

NEDO human cDNA sequencing project

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 2228)

TITLES

Sugano, S. and Suzuki, Y.

TITLES

Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp; Tel:81-3-5449-5286, Fax:81-3-5449-5416)

JOURNAL

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES

Location/Qualifiers

SOURCE

1..2228

ORGANISM

/organism="Homo sapiens"

MOL TYPE

/mol_type="mRNA"

DB XREF

/db_xref="taxon:9606"

CLONE

/clone="CBR05878"

TISSUE

/tissue_type="brain"

CLONE LIB

/clone_lib="CBR"

NOTE

/note="Cloning vector: pME18SFL3"

BASE COUNT

478 a 598 c 696 g 456 t

ORIGIN

Alignment Scores:

Pred. No.: 3..546-173 Length: 2228
 Score: 165.00 Matches: 165
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 59.57% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x AK057848 (1-2228)

QY 37 SerThrValGlyGluLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGln 56
 Db 214 TCAAGCTGTGTAGAAAGATCATGAGTCTCGGACAGAACAGAGTGGAAACGCGGACG 273
 QY 57 ArgLysValValIleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLys 76
 Db 274 CGAAGGTGTCTATCTCGAGCCAGGACAGGTTCTACAAGTCTCGACGGCAGAGCAAG 333
 QY 77 AlaLysAlaLeuLysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeu 96
 Db 334 GCCAAGCCCTTGAAGAGACGTCAATTGACCAATCCAGATGCTTTGATTAATGATTG 393
 QY 97 MetHisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAsp 116
 Db 394 ATGCACGACCTTGAAGAAATCATCGTGGAGGGCAAAACGTTGAGGTGCCGACCTTATG 453
 QY 117 PheValThrHisSerArgLeuProGluThrThrValValTyrProAlaAspValValLeu 136

```

Db      454 TTTGTGACACACTCAAGTTACAGAGACACAGTGTCTACCTCGGACGTGTTCTG 513
Qy      137 PheGluGlyIleuValPheTYSerGlnGluIleArgApwecPheHi sleuArgLeu 156
Db      514 TTTGAGGGCATCTTGTTGTTCTACAGCCAGAGATCCGGAGACATGTTCCACCTCGCGCTC 573
Qy      157 PheValAspThrAspSerAspValArgLeuSerArgArgValIleuArgAspValArgArg 176
Db      574 TTGTGGACACCGACCTCGACGTCGAGGCTGTCTGAAAGATTCTCCGGAGCGTGGCGGA 633
Qy      177 GYAArgApwecGluGlnIleleuThrGlnTYrThrPheValIysProAlaPheGlu 196
Db      634 GGGAGGGACCTGAGCAGATTTCGACCGCAGTACCACTTCCTCGAAGCCGCGCTTCGAG 693
Qy      197 GluPheCysLeuPro 201
Db      694 GAGTTCTGCGCTGCGC 708

RESULT 11
AX540411
LOCUS      AX540411          734 bp      DNA      linear      PAT 23-NOV-2002
DEFINITION Sequence 23 from Patent WO02055738.
ACCESSION  AX540411
VERSION     AX540411.1  GI:25273437
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
  1 Panzer, S.R., Lincoln, S.E., Altue, C.M., Dufour, G.E., Hillman, J.L.,
    Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Paito, A.,
    Marwaha, R., Chen, A.J., Chang, S.C., Gerstlin, E.H., Peralta, C.H.,
    David, M.H. and Lewis, S.A.
    Molecules for disease detection and treatment
    Patent: WO 02055738-A 23 18-JUL-2002;
    Incyte Genomics, Inc. (US)

FEATURES
    source
        1..734
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /note="Incyte ID No: LI:235557.12:2001JAN12"

BASE COUNT      166 a      181 c      226 g      154 t      7 others

ORIGIN
Alignment Scores:
Pred. No.:      1 28e-114      length:      734
Score:          112.00      Matches:      112
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    40.43%      Indels:      0
                Gaps:      0

US-09-896-522-2 (1-277) x AX540411 (1-734)

Qy      90 AspAlaPheAspAsnAspLeuMetHi sArgThrIleuYsaAnIleValIgluGlyLysThr 109
Db      14 GATGACCTTTGATATGATTGATGACACAGACTCTGAAGAACATCGTGAGGGCAAAACG 73
Qy      110 ValGluValProThrTYrAspPheValThrHisSerArgLeuProGluTYrThrValVal 129
Db      74 GTGAGAGTCCGACCTATGATTTGTGACACACTCAAGATTACCGAGACACCGTGGTC 133
Qy      130 TYrProAlaAspValValIleuPheGluGlyIleuValPheTYrSerGlnGluIleArg 149
Db      134 TACCTCGCGGACGTGGTCTGTTGAGGGCATCTTGTTCTTACAGCCAGAGATCCGG 153
Qy      150 AspMetPheHi sleuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
Db      194 GACATGTTCCACCTGCGCTCTTGTGAGAACCGACTCCAGCGTCAAGGCTGTCTCGAAGA 253

```

```

Qy      170 ValIeuArgAspValArgArgGlyArgAspLeuGluGlnIleleuThrGlnTYrThr 189
Db      254 GTTCTCCGGAGCGTGGCCGAGGAGGAGACCTGAGACATTTGACGAGTACACACACC 313
Qy      190 PheValIysProAlaPheGluGluPheCysLeuPro 201
Db      314 TTGTGAGCGCGGCTTCGAGAGATTCTGCTCGCCG 349

RESULT 12
AL358781/c
LOCUS      AL358781          147492 bp      DNA      linear      PRI 06-OCT-2001
DEFINITION Human DNA sequence from clone RP11-334J6 on chromosome 9, complete
            sequence.
ACCESSION  AL358781
VERSION     AL358781.19  GI:13751418
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 147492)

REFERENCE
  Corby, N.
  Direct Submision
  Submitted (06-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
  requests: clonerequest@sanger.ac.uk
  On Apr 21, 2001 this sequence version replaced gi:13396549.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submision
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one plasmid subclone or more than one M13 subclone; and the
  assembly was confirmed by restriction digest. The following
  abbreviations are used to associate primary accession numbers given
  in the feature table with their source databases: Em: EMBL; Sw:
  SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP
  database can be found at
  http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
  was generated from part of bacterial clone contigs of human
  chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
  Group. Further information can be found at
  http://www.sanger.ac.uk/HGP/Chr9
  RP11-334J6 is from the library RPCT-11.2 constructed by the group
  of Pieter de Jong. For further details see
  http://www.chori.org/bacpac/home.htm
  VECTOR: pBAC3.6
  This sequence is the entire insert of clone RP11-334J6 The true
  left end of clone RP11-40A7 is at 113870 in this sequence. The true
  right end of clone RP11-643B14 is at 63282 in this sequence.

FEATURES
    source
        1..147492
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="9"
        /clone="RP11-334J6"
        /clone_1ib="RPCT-11.2"
        /clone_1ib="RPCT-11.2"
        misc_feature
            30143..30199
            /note="Sequence from overlapping clone RP11-643B14
            (AL354855). Assembly confirmed by restriction digest."

BASE COUNT      32992 a      36290 c      37577 g      40633 t

ORIGIN
Alignment Scores:
Pred. No.:      1.68e-54      Length:      147492
Score:          60.00      Matches:      60

```

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 21.66% Indels: 0
 DB: 9 Gaps: 0

US-09-896-522-2 (1-277) x ALJ58781 (1-147492)

Qy 218 ValAlaIleAenLeuIleValIGlnHisIleGlnAspIleLeuAenGlyAspIleCysIleYs 237
 Db 129488 GTTGCACATCAACCTGATGTCGACAGACATCCAGACATTCGATGTCGATCTGCATA 129429

Qy 238 TTPHisArgIlyGlySerAsnGlyAArgSerTyrIlyAspArgThrPheSerGluProGlyAsp 257
 Db 129428 TGGACACGAGGAGGCTCCAAATGGCGGAGCTACAAAGCGAGCTTTCTGAGCCAGGGGAC 129369

Qy 258 HisProGlyMetLeuThrSerGlyIlyAspSerHisIleGlnGlySerSerSerArgProHis 277
 Db 129368 CACCTGGAGTGTGACCTCTGGCAACGATCACTTTGGAGTCACGACGACCCAC 129309

RESULT 13

MUSURKI 1810 bp mRNA linear ROD 27-MAR-1997
 LOCUS DEFINITION Mus musculus uridine kinase mRNA, partial cds.
 ACCESSION L31783
 VERSION L31783.1 GI:471980
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1810)
 AUTHORS Traut, T.W.
 TITLE The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites

JOURNAL Eur. J. Biochem. 222 (1), 9-19 (1994)
 MEDLINE 94259063
 PUBMED 8200357

REFERENCE 2 (bases 1 to 1810)
 AUTHORS Ropp, P.A. and Traut, T.W.
 TITLE Cloning and expression of a cDNA encoding uridine kinase from mouse brain

JOURNAL Arch. Biochem. Biophys. 336 (1), 105-112 (1996)
 MEDLINE 97108719
 PUBMED 8951040

FEATURES

source Location/Qualifiers
 1..1810
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 <1..783
 /BC_number="2.7.1.48"
 /standard_name="uridine-cytidine kinase"
 /codon_start=1
 /product="uridine kinase"
 /protein_id="PAB50568.1"
 /db_xref="GI:471981"
 /translation="RPPRPFLIGSVGTASGKTVCEKIMELIGONEVDRRQKVI
 LSODCFKVLTAQKAKAIKQGVNPDADPNDIMKTKIKNIVEGTVVPPYDPT
 HSRLEPVTVPADVLPFGLIVPTQEIPIRDMFLRLPVTDSVRLSRVLRDVORG
 RDLQILQYIAFVAPAEFEFCLEPTKRYADVITPRGVNVALNLIVOHIDILNGDL
 CKRRRGNGNHNKRTPEEPDHPGVATGRSHLESSRPH"
 join(37..63,412..426,451..465)
 /bound_moiety="ATP at catalytic site"
 1790..1795

CDS
 misc_binding
 polyA_signal
 BASE COUNT 411 a 482 c 512 g 405 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,25e-39 Length: 1810
 Score: 45.00 Matches: 164
 Percent Similarity: 95.35% Conservative: 0
 Best Local Similarity: 95.35% Mismatches: 4

Query Match: 16.25% Indels: 8
 DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x MUSURKI (1-1810)

Qy 67 PheTyrIlyValIleuThrIaGlnIlyAlaIlyAlaIleuIlyGlyIntYrAenPhe 86
 Db 148 TTCACAAAGTCTGACGCGCGAGAGAGCCAAAGCCACTTGAAGGACAGTCAATTTT 207

Qy 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArg-ThrLeuIlyAsnIleValGI 106
 Db 208 GACCACCCCAATCTTTGTGAATGATCGATGCACA-GACCTGAAAAACATTGTGA 266

Qy 106 uGlyIlyThrValIGlIValProThrTyrAspPheValThrIleSerArgLeuProGluTh 126
 Db 267 AGCCAAACCTGTCGAGAGTCCCTACCTATGATTTTGGACCCACTCAAGTTACAGAGAC 326

Qy 126 rThrValIalTyrProAlaAspValIleuPheGluGlyIleLeuValPheTyrSer-G 146
 Db 327 CACTGTGTCTACCCAGCTGATGTGTCTGTTCGAGGCGCATCTTGATTTCTACAC-CC 385

Qy 146 InGIuIleArgAspMetPheHisLeuArgIleuPheValAspThrAspSerAspValArgI 166
 Db 386 AGGAGATCCGGGACATGTTCCACCTGGCTCTTTGGACACAGACTGTGATTTAGGC 445

Qy 166 euseArArgValIleuArgAspValArg-ArgGlyAArgSpLeuGluGlnIleLeuThr 185
 Db 446 TGTCTGAAAGATGTTCTCCGGAGTGTCG-ACGAGAAAGGACCTGAGACAGATCTGACT 504

Qy 186 GIITyrThrThr-PheValIlyProAlaPheGluGluPheCysLeuProThrIlySerTy 205
 Db 505 CAGTACACCCG-CTTGTGAAACACAGCTTTGGAGGATTCCTGCTCCGACTAAGAGTA 563

Qy 205 rAlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGI 225
 Db 564 CGCTGACGATATATCCCTCGAGAGTGTATATATGTGGCCATCAACCTGATCGTGA 623

Qy 225 nHisIleGlnAspIleLeuAsnGlyAsp 234
 Db 624 ACAATCTCAGACATCTTCAACGGGAC 651

RESULT 14

BC025146 1959 bp mRNA linear ROD 16-APR-2003
 LOCUS DEFINITION Mus musculus uridine monophosphate kinase, mRNA (cDNA clone
 MGC:36231 IMAGE:4913412), complete cds.
 ACCESSION BC025146
 VERSION BC025146.1 GI:19263563
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1959)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sableiron, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, A.C., Grimwood, J., Schmitz, J., Myers, R.M., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length

JOURNAL human and mouse cDNA sequences
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 22388257
 REFERENCE 12477932
 AUTHORS 2 (bases 1 to 1959)
 TITLE Strausberg, R.
 JOURNAL Direct Submission
 Submitted (05-MAR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Guaratine, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati,
 A.N., Gibbs, R.A.

FEATURES
 source
 1. 1959
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:36231 IMAGE:4913412"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_id="NCI_CGAP_SG2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 1959
 /gene="Umpk"
 /db_xref="LocusID:22245"
 /db_xref="MGI:98904"
 57. 890
 /codon_start=1
 /product="Umpk protein"
 /protein_id="AAH25146.1"
 /db_xref="GI:19263564"
 /db_xref="LocusID:22245"
 /translation="MASAGGSESAPADPPFLIGSGTASGKSTCEKIM
 ELIGQVDRRORLVILODFKYLAKKAKGPNPDAFNDLMEKTLX
 NIVEGTVRPTDFTVTHSLPRTVVPDVLFEILVFTYQERIDMHLAFDPT
 DSVRLSRRLRVQGRDLEQILITOTATKAFBEFCPTKYADAVIIPKVDNVP
 AINLIVOHIDILNGDICRHRGCGPNRHKRTFPPEGDHPVLATGKRSLSSRP
 H"

gene

CDS

BASE COUNT 466 a 511 c 561 g 421 t
 ORIGIN

Alignment Scores:

Score: 1.34e-39 Length: 1959
 Percent Similarity: 45.00 Matches: 164
 Best Local Similarity: 95.35% Conservative: 0
 Query Match: 95.35% Mismatches: 4
 DB: 16.25% Indels: 8
 Gaps: 0

US-09-896-522-2 (1-277) x BC025146 (1-1959)

Qy 67 PhetyrlyseValleuthralaGlunllyAlalysalaleuysgIyGlnTyranpHe 86

|||||
 Db 255 TTCTACAGAGTTCTGACGGCCGAGAGAGCCAGGCTTTGAAGGAGACGATCAATTTT 314
 Qy 87 AsphIsProAspAlaPheAspAsnAspLeuMethIaArg-ThrLeuIysAsnIleValGI 106
 Db 315 GACCACCCAGATCTTTGTATATGATCTGATCACA-A-GACCTTGAAAAATTGTGA 373
 Qy 106 uGIyVsrThrValaGlunValProThrTyraAspPheValThrHisSerArgLeuProGIuTh 126
 Db 374 AGGCAAACTGTGAGAGTCCCTACCTATGATTTTGGACCCACTCAAGGTTACCAAGAC 433
 Qy 126 rThrValValTyProAlaAspValValleuPheGIuIleLeuValPheTyser-G 146
 Db 434 CACTGTGCTTACCAAGCTGATGTGTTCTGTTCGAGGGATCTTGGTATTCACAC-CC 492
 Qy 146 InGIuIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgL 166
 Db 493 AGGAGATCCGGGACATGTTCCACCTCGCTCTTTGTGACACAGACTCGATGTTAGGC 552
 Qy 166 euserArgArgValleuArgAspValArg-ArgGIyArgAspLeuGIuIleLeuThr 185
 Db 553 TGTTCTGAAGATGTTCTCCGGATGTCA-ACGAGGAAGGACCTGAGAGATCTGACT 611
 Qy 186 GlnTyrThrThr-PheValIysProAlaPheGIuIlePheCysLeuProThrTyrsTy 205
 Db 612 CAGTACACCCGCTTTGTGAAACCAAGCTTTGAGAGCTTGCCTCCGACTAAGAGTA 670
 Qy 205 rAlaAspValIleIleProArgGIyValAspAsnMetValAlaIleAsnLeuIleValGI 225
 Db 671 CGCTGACGATCATCTCTCGAGAGTTGATATATGTGGCCATCAACCTGATCTGCA 730
 Qy 225 nHisIleGlnAspIleLeuAsnGIyAsp 234
 Db 731 ACAATCCAGACATCTCTCAACGGGAC 758

RESULT 15

AC078885/c

LOCUS

254993 bp DNA linear HTG 15-MAY-2002
 Mus musculus chromosome 6 clone RP23-117723 strain C57BL6/J,
 WORKING DRAFT SEQUENCE, 44 unordered pieces.

DEFINITION

ACCESSION

AC078885

VERSION

AC078885.8 GI:18376842

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Ferreira, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.

TITLE

High Throughput Mouse Sequencing

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 254993)

AUTHORS

Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
 Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Ferreira, A.,
 Gordon, M., Goltz, J.S. and Kucherlapati, R.

TITLE

Direct Submision

JOURNAL

Submitted (08-AUG-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Jan 26, 2002 this sequence version replaced gi:14488282.

COMMENT

-----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site: <http://www.hpcg.org/Sequence/mouse.html>
 Contact: hpcg@med.nyu.edu
 -----Summary Statistics
 Center project name: AAP
 Sequencing vector: pUC18, L08752
 Chemistry: Dye-terminator Big Dye, 100%
 *Consensus quality: 247096 at least Q20
 *Consensus quality: 242518 at least Q30

* Consensus quality: 235589 at least Q40
* Estimated insert size: agarose-fp - N/A
* Estimated insert size: 254133 - sum-of-contigs
Quality coverage: agarose-fp - N/A
Quality coverage: 6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 27415: contig of 27415 bp in length
* 27416 27435: gap of unknown length
* 27436 57817: contig of 30382 bp in length
* 57818 57837: gap of unknown length
* 57838 81637: contig of 23800 bp in length
* 81638 81657: gap of unknown length
* 81658 102155: contig of 20498 bp in length
* 102156 102175: gap of unknown length
* 102176 121200: contig of 18025 bp in length
* 121201 121220: gap of unknown length
* 121221 135073: contig of 13853 bp in length
* 135074 135093: gap of unknown length
* 135094 156689: contig of 21596 bp in length
* 156690 156709: gap of unknown length
* 156710 166185: contig of 9476 bp in length
* 166186 166205: gap of unknown length
* 166206 178798: contig of 12593 bp in length
* 178799 178818: gap of unknown length
* 178819 187247: contig of 8429 bp in length
* 187248 187267: gap of unknown length
* 187268 192512: contig of 5225 bp in length
* 192513 192519: gap of unknown length
* 192520 200084: contig of 7572 bp in length
* 200085 200104: gap of unknown length
* 200105 207054: contig of 6950 bp in length
* 207055 207074: gap of unknown length
* 207075 212944: contig of 5870 bp in length
* 212945 219664: gap of unknown length
* 219665 219468: contig of 6504 bp in length
* 219469 219488: gap of unknown length
* 219489 220369: contig of 881 bp in length
* 220370 220389: gap of unknown length
* 220390 221164: contig of 775 bp in length
* 221165 221184: gap of unknown length
* 221185 222796: contig of 1612 bp in length
* 222797 222816: gap of unknown length
* 222817 224762: contig of 1946 bp in length
* 224763 224782: gap of unknown length
* 224783 225999: contig of 1217 bp in length
* 226000 226019: gap of unknown length
* 226020 227409: contig of 1390 bp in length
* 227410 227429: gap of unknown length
* 227430 230145: contig of 2716 bp in length
* 230146 230165: gap of unknown length
* 230166 231842: contig of 1677 bp in length
* 231843 231862: gap of unknown length
* 231863 232786: contig of 924 bp in length
* 232787 232806: gap of unknown length
* 232807 233137: contig of 331 bp in length
* 233138 233157: gap of unknown length
* 233158 234519: contig of 1362 bp in length
* 234520 234539: gap of unknown length
* 234540 236416: contig of 1877 bp in length
* 236417 236436: gap of unknown length
* 236437 237967: contig of 1531 bp in length
* 237968 237987: gap of unknown length
* 237988 239220: contig of 1233 bp in length
* 239221 239240: gap of unknown length
* 239241 239267: contig of 127 bp in length

* 239368 239387: gap of unknown length
* 239388 239399: contig of 572 bp in length
* 239960 239979: gap of unknown length
* 239980 241016: contig of 1037 bp in length
* 241017 241036: gap of unknown length
* 241037 242391: contig of 1355 bp in length
* 242392 242411: gap of unknown length
* 242412 243823: contig of 1412 bp in length
* 243824 243843: gap of unknown length
* 243844 245227: contig of 1384 bp in length
* 245228 245247: gap of unknown length
* 245248 246334: contig of 1087 bp in length
* 246335 246354: gap of unknown length
* 246355 247451: contig of 1097 bp in length
* 247452 247471: gap of unknown length
* 247472 248780: contig of 1309 bp in length
* 248781 248800: gap of unknown length
* 248801 250041: contig of 1241 bp in length
* 250042 250061: gap of unknown length
* 250062 251173: contig of 1112 bp in length
* 251174 251193: gap of unknown length
* 251194 252713: contig of 1520 bp in length
* 252714 252733: gap of unknown length
* 252734 254014: contig of 1281 bp in length
* 254015 254034: gap of unknown length
* 254035 254176: contig of 142 bp in length
* 254177 254196: gap of unknown length
* 254197 254993: contig of 797 bp in length.
Location/Qualifiers
1. .254993
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-117I23"
/sex="male"
1. .27415
/note="assembly_name:Contig138"
27436. .57817
/note="assembly_name:Contig137"
clone_end:586
vector_side:left"
57838. .81637
/note="assembly_name:Contig136"
81658. .102155
/note="assembly_name:Contig135"
102176. .121200
/note="assembly_name:Contig134"
121221. .135073
/note="assembly_name:Contig133"
135094. .156689
/note="assembly_name:Contig132"
156710. .166185
/note="assembly_name:Contig131"
clone_end:77
vector_side:left"
166206. .178798
/note="assembly_name:Contig130"
178819. .187247
/note="assembly_name:Contig129"
187268. .192492
/note="assembly_name:Contig128"
192513. .200084
/note="assembly_name:Contig127"
200105. .207054
/note="assembly_name:Contig126"
207075. .212944
/note="assembly_name:Contig125"
212965. .219468
/note="assembly_name:Contig124"
219489. .220369
/note="assembly_name:Contig123"

```

misc_feature      220390..221164
                  /note="assembly_name:Contig122"
misc_feature      221185..222796
                  /note="assembly_name:Contig121"
misc_feature      222817..224762
                  /note="assembly_name:Contig120"
misc_feature      224783..225999
                  /note="assembly_name:Contig119"
misc_feature      226020..227409
                  /note="assembly_name:Contig118"
misc_feature      227430..230145
                  /note="assembly_name:Contig117"
misc_feature      230166..231842
                  /note="assembly_name:Contig116"
misc_feature      231863..232786
                  /note="assembly_name:Contig115"
misc_feature      232807..233137
                  /note="assembly_name:Contig114"

```

Alignment Scores:

```

Pred. No.:      1.44e-37      Length:      254993
Score:          45.00          Matches:      77
Percent Similarity: 97.47%      Conservative: 0
Best Local Similarity: 97.47%      Mismatches: 1
Query Match:    16.25%          Indels:      2
DB:             2              Gaps:          0

```

US-09-896-522-2 (1-277) x AC078885 (1-254993)

```

QY      67      PheTYrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlnTYrAsnPhe 86
          |||||
Db      156571 TTCTACAGGTTCTGACGGCGAGCAGAGCCAGGCTTGAAGGACAGTACAATTTT 156512
          |||||

QY      87      AspHisProAspAlaPheAspAsnAspLeuMetHisArg-ThrLeuLysAsnIleValG1 106
          |||||
Db      156511 GACCACCACGATGCTTTGATTAATGATGATGACACAA-GACCCGAAAAAACATTGTTGA 156453
          |||||

QY      106     uG1yLysThrValGluValProThrTYrAspPheValThrHisSerArgLeuProGluThr 126
          |||||
Db      156452 AGGCAAAACTGTGCGAGGTCCTACTATGATTTTGTGACCCCACTCAAGGTTACCGAGAC 156393
          |||||

QY      126     rThrValValTYrProAlaAspValValLeuPheGluGlyIleLeuValPheTYr 144
          |||||
Db      156392 CACTGTGTGTACCACTGATGTGTCTGTTCGAGGGCATCTTGATTTCTAC 156338
          |||||

```

Search completed: November 25, 2003, 11:28:57
 Job time : 4109 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 25, 2003, 08:11:01 ; Search time 2545 Seconds
(without alignments)
2645.320 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 1450
Sequence: 1 MASAGGDECPAPBEADRP.....HPGMLTSGKSHLESSSRPH 277

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DB=x1h
-Q=/cgr2_1/USPRO.spool/US09896522/rumac_21112003_184354_5011/app_query.fasta_1.455
-DB=EST -QMT=fastlap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR=US09896522.@CGN_1.1_2810.@rumac_21112003_184354_5011 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : EST:
1: em_estba:*
2: em_estbm:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1401	96.6	1118	12	BM467984 AGENCOURT
2	1346	92.8	1201	13	BX394295 BX394295
3	1332	91.9	1114	13	BX343101 BX343101
4	1327	91.5	998	13	BX400889 BX400889
5	1316	90.8	916	12	BG826894 602750978
6	1314	90.6	860	10	BG491358 602535642
7	1313	90.6	1051	12	BM459108 AGENCOURT
8	1311.5	90.4	869	10	BG491384 602535670
9	1300.5	89.7	848	12	B1258532 602972340
10	1278	88.1	1044	12	BM917506 AGENCOURT
11	1277	88.1	956	13	BQ719741 AGENCOURT
12	1274	87.9	943	13	BQ511803 AGENCOURT
13	1270	87.6	1036	13	BQ072501 AGENCOURT
14	1260	86.9	961	12	B1416102 602987387
15	1259	86.8	922	10	BG390519 602416183
16	1257	86.7	922	13	BQ935919 AGENCOURT
17	1246	85.9	827	12	B1261258 602969110
18	1238	85.4	810	12	B1753007 603025788
19	1232	85.0	779	9	AU131406 AGENCOURT
20	1226	84.6	819	12	B1910015 603068126
21	1224	84.4	725	12	B1256928 602975534
22	1219.5	84.1	959	12	B1770572 603060368
23	1215	83.8	796	9	AU141725 AGENCOURT
24	1199	82.7	889	13	BQ885655 AGENCOURT
25	1195	82.4	898	14	CD514811 AGENCOURT
26	1188.5	82.0	819	12	B1770639 603060460
27	1183	81.6	822	12	BM948090 AGENCOURT
28	1168.5	80.6	910	13	BQ437115 AGENCOURT
29	1132	78.1	952	10	BG163897 AGENCOURT
30	1119	77.2	812	12	BG974826 602842705
31	1115.5	76.9	1055	12	BM545603 AGENCOURT
32	1113.5	76.8	875	13	BQ714585 AGENCOURT
33	1107	76.3	772	10	BF346907 602021736
34	1105	76.2	730	12	BG770518 602734236
35	1096	75.6	768	13	BQ614140 AGENCOURT
36	1083.5	74.7	876	13	BQ292179 AGENCOURT
37	1074.5	74.1	753	9	AU122008 AGENCOURT
38	1069	73.7	778	14	CD101693 AGENCOURT
39	1068.5	73.7	734	14	CA386173 AGENCOURT
40	1058	72.7	1039	12	BM803129 AGENCOURT
41	1054	72.0	1095	12	BM802939 AGENCOURT
42	1041.5	71.8	912	12	B1412715 602987440
43	1030	71.0	866	12	B1526272 602926190
44	1030	71.0	694	12	B1906603 603064270
45	1009.5	69.6	863	12	B1085319 602870377

ALIGNMENTS

RESULT 1
LOCUS BM467984 1118 bp mRNA EST 05-FEB-2002
DEFINITION AGENCOURT_6437937 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532885
5', mRNA sequence.
ACCESSION BM467984
VERSION BM467984.1 GI:18517026
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1118)

AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnll.gov>
 Plate: LM142216 row: n column: 22
 High quality sequence: 689.

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="5532885"
/risue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="N1H MGC 71"
/notes="Organ: uterus; Vector: pCMV-Sport6; Site 1: NotCl; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 2.1 Kb."

```

Alignment Scores:

Pred. No.:	4.72e-157	Length:	1118
Score:	1401.00	Matches:	271
Percent Similarity:	98.19%	Conservative:	1
Best Local Similarity:	97.83%	Mismatches:	4
Query Match:	96.62%	Indels:	
DB:	12	Gaps:	0

US-09-896-522-2 (1-277) X BM467984 (1-1118)

Qy		MetAlserAglIyGlnAspYsgYsIuseRProAlaProGluAlaAspArgProHis	20
Db	47	ATGGCTTCGGCGGAGGCGAGACCTGCGAGACCCCGCGGAGCGACCGTCCGAC	106
Qy	21	GlnArpProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyYlySerThrValCys	40
Db	107	CAGGCGCCCTTCGTGATAGGCGGTGACGGCGGCACTGCCAGGGGGAAGTCACCTGTGT	166
Qy	41	GlnIysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal	60
Db	167	GAGAAAGTCAATGCAATGCTGGGACAGAACGAGGTGGAAACGGCGCACGGGAAGTGTGTC	226
Qy	61	IleIuseSerGlnAspArgPheTyrIysValIleAsnThrAlaGluGlnIlyValAlaVal	80
Db	227	ATCTGAGCCAGACAGACAGTTCThCAAGGCTCTGACGGACAGACAGAACGACCAAGGCTTGG	286
Qy	81	LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
Db	287	AAAGGACAGTACAAATTTTGACCATCCAGATGCCCTTGTAATGATTTGATGACAGGACT	346
Qy	101	LeuIlyAsnIleValGluGlyIlyThrValAlaGluValProThrTyrAspPheValThHis	120
Db	347	CTGAAAGAACATCGTGAGGGCGAAACAGGGGAGGGCGGACCTTGATTTGTTGACACAC	406
Qy	121	SerArgLeuProGluThrThrValValTyrProAlaAspValIleLeuPheGluGlyIle	140
Db	407	TCAGAGTTACCAAGACCAACGGGTGGTCTACCTCGGACGATGGTTCTGTTTGAGAGCATC	466
Qy	141	IleuValPheTyrSerGlnGluIleArgAspMetPheHisIleArgIlePheValAspThr	160
Db	467	TTGGTGTTCACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTCTTGCTGGACACC	526
Qy	161	AspSerAspValArgLeuSerArgArgValIleuArgAspValArgArgGlyArgAspLeu	180

Db	527	GACGCCAGCGCAGAGCTGTCTCGAAGAGTTCTCCGGAGCGCGCCGAGGAGGACCTG	586
Qy	181	GlnGlnIleLeuThrGlnIlyrThrThrPheAllylsProAlaPheGlnGlnPheCysIleu	200
Db	587	GAGAGAGATTGACGCGAGTACACCACTTCGTGAAGCCGCGCTTCGAGAAATTCTGGCTG	646
Qy	201	ProThrIlysIlyrAlaAspValIleIleProArgGlyValaAspAsmMetValAlaIle	220
Db	647	CCGACAAAGAAAGTATGCCAGATGTATCATCTCCACGAGGAGTGGCAATATGTTGGCATTC	706
Qy	221	AsnLeuIleValGlnHisIleIleGlnAspIleLeuAsnGlyAspIleCysIlyrThrPheArg	240
Db	707	AACTGTATCTGGCGACATCTCAGAGACATTTCTATGTGTGCATCTGCAAAATGGACCGA	766
Qy	241	GlyIlySerAsnGlyIlyrSerTyrIlyAsrThrPheSerGlnProGlyAspHisAspGly	260
Db	767	GGAAGGTTCAAATGGCGGAGCTACAAAGGAGCCTTTTCTGAGCAGGGGAGCACCTCTG	826
Qy	261	-MetLeuThrSerGlyIlyrAsrSerHisLeuGlnSerSerArgPro	276
Db	827	GATCTGATCCTCTGTGCAACGGTTCACATTTGAGATTCAGCGAGACCC	875

RESULT 2	
BX394295	
LOCUS	BX394295
DEFINITION	1201 bp mRNA linear EST 13-MAY-2003
ACCESSION	BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
VERSION	CDNA clone CS0D0013YE13 5'-PRIME, mRNA sequence.
KEYWORDS	BX394295 GI:30624219
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1. (bases 1 to 1201)
Li, W. B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
[http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC013AC07Qp1&cluster=4968.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC013AC07Qp1&cluster=4968.r). Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DC013AC07Qp1.

FEATURES

SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="txon:9606"
/clone="OS0DC013YE13"
/feature_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_1ib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMSPORT 6 vector. Library was normalized."

```

Alignment Scores:

Pred. No.:	2,09e-150	Length:	1201
Score:	1346.00	Matches:	270
Percent Similarity:	98.55%	Conservative:	
Best local Similarity:	97.83%	Mismatches:	4
Query Match:	92.83%	Indels:	
DB:	13	Gaps:	0


```

Db      638 GACTCCAGCTCAGGCTGCTCTCGAAGATTCTCCGGAGCTGCCCGAGGGAGGACTTG 697
Qy      181 GlnGlnIleLeuThrglnTyThrThrPheValIysProAlaPheGlnGluPheCysLeu 200
Db      698 GAGCAGATTCTGACGCAATACACCACTTGCTGTAAGCCGGCTTCGAGGAGTTCTGCTG 757
Qy      201 ProThrIlySerTyTyAlaAspValIleIleProArgIlyValAlaPheMetValAlaIle 220
Db      758 CCGCAAGAAGATGATGCGATGATGATCCACAGAGATGAGCAATATGTTGCCATC 817
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIlySTpHisArg 240
Db      818 AACCTGATCTGCGACGACATCCAGACATTTCTGATGATGACATCTGCAATGACACCGA 877
Qy      241 GlyIlySerAsnGlyArgSerTyTyAlaArgThrPheSerGluProGlyAspHisProGly 260
Db      878 GGAAGGTCATGAGCGGAC-TACAGCGGACCTTTTCTGAGCCAGGGGA-CACCTGGG 935
Qy      261 MetLeuThrSerGlyIlyArgSerHisLeuGluSerSerArgPro 276
Db      936 ATGCTGACTCTGCGAAACGGTCACTTGGAGTCAGACAGACACCA 983

RESULT 4
LOCUS   BX400889 998 bp mRNA linear EST 13-MAY-2003
DEFINITION BX400889 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK003YL24 5'-PRIME, mRNA sequence.
ACCESSION BX400889
VERSION   BX400889.1 GI:30622359
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS L.J.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE    Full-length cDNA libraries and normalization
JOURNAL  Unpublished
COMMENT  Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK003DF12QPI&cluster=4968.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DK003DF12QPI.
FEATURES
source
location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK003YL24"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 223 a 263 c 314 g 186 t 12 others
ORIGIN
Alignment Scores:
Pred. No.: 2,99e-148 Length: 998
Score: 1327.00 Matches: 271
Percent Similarity: 97.83% Conservative: 0
Best Local Similarity: 97.83% Mismatches: 6
Query Match: 91.52% Indels: 4

```

```

DB: 13 Gaps: 0
US-09-896-522-2 (1-277) x BX400889 (1-998)
Qy      1 MetAlaSerAlaGlyIlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
Db      67 ATGCTTCGCGCGGAGGCGGAAGCTGCGAGAGCCCGCCGCGGA-SCGNAACGTCGCGAC 125
Qy      21 GlnArgProPheLeuIleGlyValSerGlyIlyThrAlaSerGlyIlySerThrValCys 40
Db      126 CAGCGNCCCTCTCTGATPAGGGGTGACGGCGGACGACCGAGCGGAAGTCGACCCGTGT 185
Qy      41 GlnIlyIleMetGlnIleLeuLeuGlnIlyValGlnIlyValGlnIlyValGlnIlyVal 60
Db      186 GAGAGATCTAGAGATGCTGCTGGGACGAAACGAGTGAACAGCGGACGCGGAAGTGTCT 245
Qy      61 ILeuSerGlnAspArgPheTyTyAlaValLeuThrAlaGlnIlyValAlaIlyValAlaLeu 80
Db      246 ATCTGAGCGAGGACAGGTTCTCAAGGTCCTGACCGCAGAGACGAAAGCCAGGCTTG 305
Qy      81 LysGlyGlnTyTranPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db      306 AAAGGACAGTACAAATTTGACATCCAGATGCTTGTGATTAATGATTTGATGACAGAGCT 365
Qy      101 LeuIlyAsnIleValGlnIlyIlySerThrValGlnIlyValProThrTyTranPheValThrHis 120
Db      366 CTGAAGAACATCTGTGAGAGGCGAAGCGGTGAGGTGCCAGCTTATATTTTGTGACACAC 425
Qy      121 SerArgLeuProGluThrThrValValIlyProAlaAspValIlyLeuPheGlnIlyIle 140
Db      426 TCAAGGTTCACAGAGACACAGGTTGCTACCCGCGAGCGTGGTCTGTTTGGGGCATC 485
Qy      141 LeuValPheTyTranSerGlnIlyIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db      486 TTGGTTTCTACAGCGCAGAGATCCCGGACATGTTCCACTGGCCCTCTTCGTGGACACC 545
Qy      161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgIlyArgAspLeu 180
Db      546 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGCGCCGAGGAGGACCTG 605
Qy      181 GlnGlnIleLeuThrglnTyThrThrPheValIysProAlaPheGlnGluPheCysLeu 200
Db      606 GACGAGATTCTGACGCA-TACACCACTTCGTGAACCGGCTTCGAGGAGTTCTGCTG 664
Qy      201 ProThrIlySerTyTyAlaAspValIleIleProArgIlyValAlaPheMetValAlaIle 220
Db      665 CCACCAAGAAGATATCCGATGATCATCCACGAGAGTGAACAATATGTTGCCATC 724
Qy      221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIlySTpHisArg 240
Db      725 AACCTGATCTGCGACGACATCCAGACATTTCTGATGATGATGATCTTGCMAATGACACGA 784
Qy      241 GlyIlySerAsnGlyArgSerTyTyAlaArgThrPheSerGluProGlyAspHisProGly 260
Db      785 GAGGAGTCCATGAGCGGAC-TACAGCGGACCTTTTCTGAGCCAGGGGA-CACCTGGG 842
Qy      261 MetLeuThrSerGlyIlyArgSerHisLeuGluSerSerArgProHis 277
Db      843 ATCTGACCTCTGCGAAACGGTCACTTGGATTCAGACGACGACGACACCCAC 893

RESULT 5
LOCUS   BG826894 916 bp mRNA linear EST 22-MAY-2001
DEFINITION BG826894 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
mRNA sequence.
ACCESSION BG826894.1 GI:14174481
VERSION   BG826894.1
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 916)

```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT

Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLCM1802 row: k column: 20
 High quality sequence stop: 843.
 Location/Qualifiers

FEATURES

source

1. 916
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4903819"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 17"
 /note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 203 a 239 c 309 g 165 t

ALIGNMENT SCORES:

Pred. No.: 5,47e-147 Length: 916
 Score: 1316.00 Matches: 270
 Percent Similarity: 96.47% Conservative: 3
 Best Local Similarity: 95.41% Mismatches: 7
 Query Match: 90.76% Indels: 3
 DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x BG826894 (1-916)

QY 1 MetAlSeRaLaGlyGluAaPcyGluSerProAlaProGluAlaAspArgProHis 20
 DB 80 ATGGCTTCCGGCGGAGCGGAGAGCTGGAGAGCCCGCGCGGAGCGCGCGCAC 139
 QY 21 GlnArgProPheLeuIleGlyValSerGlyValThrAlaSerGlyLysSerThrValCys 40
 DB 140 CAGCGGCCCTTCTGATAGGGGTGAGGGCGGCGCATGCCAGCGGAGAGTGCAGTGTG 199
 QY 41 GlnLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysVal 60
 DB 200 GAGAAAGTATGATGAGTTGCTGGAGAGAGAGAGTGGAGCAGCGGCGGAGAGTGTG 259
 QY 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaValLeu 80
 DB 260 ATCTTGAAGCAGACGAGCTTCTCAAGGTCTTCAAGCGCAGAGCAGAGCGCAAGGCTTG 319
 QY 81 LysGlyGlyTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 DB 320 AAGGAGACGTACATTTTGAATCCATCCAGATGCTTTGATATATATTTGATGACAGGACT 379
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 DB 380 CTGAGAGACATCTGTGAGGGCGCAAAACGCTGAGAGTGGCGCATATGATTTTGTGACACAC 439
 QY 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyLys 140
 DB 440 TCAAGGTTCACAGAGCAGAGGTGGTCTACCTTGGCGAGCTGTGTTCTGTTTGAAGGCGATC 499
 QY 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160

DB 500 TTGGTGTTCACAGCAGAGATCCGGAGCATGTTCCACTCGCTTCTGTGACACC 559
 QY 161 AspSerAspValArgLeuSerArgValLeuArgAspValArgArgLysArgAspLeu 180
 DB 560 GACTCCGACGTAGAGGTCTCTGAAAGAGTCTCCGGAGAGTCCGCGAGGAGGAGCTTG 619
 QY 181 GlnGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 200
 DB 620 GAGCAGATTCACGACGACACACACTTCGTGAAGCCGCGCTTCAGAGAGTTCTGCTTG 679
 QY 201 ProThrLysLysTyrAlaAspValIleLeuProArgLysValAspAsnMetValAla1 220
 DB 680 CCGACAAAGATATGCGCATGTGATCATCCACAGAGAGTGAGCAATATGTTGTCAT 739
 QY 220 eAnLeuIleValGlnHisIle-GlnAspIle-LeuAsnGlyAspIleCysLysThrHis 239
 DB 740 CAACCTGATCGTGACGACATTCACAGCATTCCTGAATGTGTGACATCTGCAAAATGGCAC 799
 QY 240 ArgGlyGlySerAsn-GlyArgSerTyrLys-ArgThr-PheSerGluProGluAspHis 258
 DB 800 CAGAGAGGTCCATAGGGCGGAGCTACAGCGGAGCTTTTCTGAGACAGGGAGAAC 859
 QY 259 Pro-GlyMetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgPro 276
 DB 860 CCTGGGAATGCTGACCTTCTGCAAAAGGTCCCATTTGAGTCCAGCAGCAGACCC 914

RESULT 6
 BG491358 860 bp mRNA linear EST 27-MAR-2001
 LOCUS BG491358 60235642F1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684738 5',
 DEFINITION mRNA sequence.

ACCESSION BG491358
 VERSION BG491358.1 GI:13452870
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 860)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

TITLE

AUTHORS

COMMENT

Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTFCDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLCM1493 row: c column: 11
 High quality sequence stop: 844.
 Location/Qualifiers

FEATURES

source

1. 860

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684738"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 195 a 223 c 288 g 154 t
 ORIGIN

Alignment Scores:

Pred. No.: 8,63e-147 Length: 860
 Score: 1314.00 Matches: 262
 Percent Similarity: 98.14% Conservative: 2
 Best Local Similarity: 97.40% Mismatches: 2
 Query Match: 90.62% Indels: 4
 DB: 10 Gaps: 0

US-09-896-522-2 (1-277) x BM459108 (1-860)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 52 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGGCCCCCGCGCGGCGGACCGTCCAC 111
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 112 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCACTGCGACGGGAGAGTCCAGCTGTGT 171
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
 Db 172 GAGAAAGATCATGAGTGTCTGGACAGAACGAGTGGAAACGCGGACGGAAGTGGTC 231
 QY 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 232 ATCTGAGCCAGGACAGGTTCTACAGGCTCTGACGCGACAGGAGGCGCAAGCCTTG 291
 QY 81 LysGlyGlnIleAsnProPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
 Db 292 AAAGAGACATGATTTTGTACCATCCAGATCCCTTGTATTAATGATTTGATGACAGACT 351
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyAspPheValThrHis 120
 Db 352 CTGAAGAACATCGTGAAGGCGAAACGGTGAAGTGGCGACTATGATTTTGTACACAC 411
 QY 121 SerArgLeuProGluThrThrValValTyProAlaAspValValLeuPheGluGlyIle 140
 Db 412 TCAGGTTACAGAGACACACGCTGTCTACCTCGGACGCTGTTCTGTTGAGGCGATC 471
 QY 141 LeuValPheTySerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 472 TTGGTGTCTTACAGCCAGAGATCGGAGCATGTTCCACCTGCGCCCTTCTGTGAGAC 531
 QY 161 AspSerAspValArgLeuSerArgArg-ValLeuArgAspValArgArgGlyArgAspLe 180
 Db 532 GACTCCGACGTCAGGCTGTCTCGAAAGATTCTCGGAGCGTGGCGCGGAGGAGGACCT 591
 QY 180 uGluGlnIleLeuThrGlnIleThrThrPheValLysProAlaPheGluGluPheCysLe 200
 Db 592 GGAACAGATTCTGACGAGATACACACCTTCTGTAAACCGGCTTCAGAGGTTCTGCT 651
 QY 200 uProThrLysIleTyAlaAspValIleIleProArgGlyValAspAsn-MetValAlaI 220
 Db 652 GCCGACAAAGAGATGCGCATGATCATCCCGAGAGTGAACAATTTGGTTTCCA 711
 QY 220 LeuLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHis- 239
 Db 712 TCAACCTGATCGTGCACACATCCAGACATTTGAAATGGTGAATCATGCAAAATGGCAC 771
 QY 240 ArgGlyLysSerAsnGlyArgSerTyLysArgThrPheSerGluProGlyAspHisPro 259
 Db 772 CGAGGAGTCCCATGAGCGGAGCTAGCAAGGACCTTTCT-GAGCGAGGAGGACCACT 830
 QY 260 GlyMetLeuThrSerGlyLys 266
 Db 831 GGAATGCTGACCTTGAGGCAA 851

RESULT 7
 BM459108 1051 bp MRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT_6414543 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557433
 DEFINITION 5', mRNA sequence.
 ACCESSION BM459108

VERSION BM459108.1 GI:18508148
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1051)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM12278 row: m column: 18
 High quality sequence stop: 542.
 Location/Qualifiers
 1..1051
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557433"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 237 a 264 c 312 g 221 t 17 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.55e-146 Length: 1051
 Score: 1313.00 Matches: 258
 Percent Similarity: 94.57% Conservative: 3
 Best Local Similarity: 93.48% Mismatches: 13
 Query Match: 90.55% Indels: 2
 DB: 12 Gaps: 0

US-09-896-522-2 (1-277) x BM459108 (1-1051)

QY 1 MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHis 20
 Db 6 ATGGCTTCGGCGGAGGCGGAAGACTGCGAGGCCCCCGCGGAGGCGGACCGTCCGAC 65
 QY 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 66 CAGCGGCCCTTCCTGATAGGGGTGAGCGCGCACTGCGACGGGAGAGTCCAGCTGTGT 125
 QY 41 GluLysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgValVal 60
 Db 126 GAAAGATCATGAGTGTCTGGGACAGAACGAGTGGAAACGCGGACGGAAGTGGTC 185
 QY 61 IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
 Db 186 ATCTGAGCCAGGACAGGTTCTTACAGGCTCTGACGCGACAGACAGAAAGGCGCTTG 245
 QY 81 LysGlyGlnIleTyAsnProPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr 100
 Db 246 AAAGAGACATGATTTTGTACCATCCAGATGCTTGTATTAATGATTTGATGACAGACT 305
 QY 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyAspPheValThrHis 120
 Db 306 CTGAAGAACATCGTGAAGGCGAAACGGTGAAGTGGCGACTATGATTTTGTGACACAC 365
 QY 121 SerArgLeuProGluThrThrValValTyProAlaAspValValLeuPheGluGlyIle 140

Db 366 TCAAGGTTACGAGACACGAGTGTCTTACCTTCGCGACGCTGTCTTGTGAGGGCATC 425
 Qy 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 426 TTGGGTGTTCAAGTCAGGAGATCCGGGACATGTTCCACCTGGCCCTTCGTGAGACACC 485
 Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 486 GATCCGACGTCAGAGATGCTCCAGAGGTTCTCCGGACGCTGCTCAGGAGGAGACCTG 545
 Qy 181 GluGlnIleLeuThrGlnTyrThrThrPheValIleProAlaPheGluGluPheCysLeu 200
 Db 546 GAGCAGATTCTGACGCGATCACCACTTCGTGAGTCCGCTCTTCGAGAGTCTGCTG 605
 Qy 201 ProThrIleValTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 Db 606 CCGACAAAGAGATGTCGATGATCATCCACGAGAGTGGACATTAATGCTGACATC 665
 Qy 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 Db 666 AACCTGATCTGCGACATCCAGACATCTGAATGTGACATCTCCAAATGTGACCT 725
 Qy 241 GlyGlySerAsnGlyArgSerTyrLys-ArgThrPheSerGluProGlyAspHis-Prog 260
 Db 726 AGAGGTCATGGGCGGAGCTCAAGTCCGACCTTTTTCGAGTCAAGGAGCACCCCTG 785
 Qy 260 LysMetLeuThrSerGlyLysArgSerHisLeuGluIleSerSer 274
 Db 786 GGATGCTGACCTCGAAGAACGGTCTCTTTGGAGTTCAGT 829
 RESULT 8
 LOCUS BG491384 869 bp mRNA linear EST 27-MAR-2001
 DEFINITION 602535670F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
 mRNA sequence.
 ACCESSION BG491384
 VERSION BG491384.1 GI:13452896
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 869)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: csaps-r@mail.nih.gov
 Tissue Procurement: DCTD/BTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1. 869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_41"
 /note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a

BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,75e-146 Length: 869
 Score: 1311.50 Matches: 264
 Percent Similarity: 96.70% Conservative: 0
 Best local Similarity: 96.70% Mismatches: 4
 Query Match: 90.45% Indels: 5
 DB: 10 Gaps: 1
 US-09-896-522-2 (1-277) x BG491384 (1-869)
 Qy 1 MetAlaSerAlaGlyGlyGluAspCysGlySerProAlaProGluAlaAspArgProHis 20
 Db 52 ATGGCTTCGCGGAGGAGCAAGCTGCGAGAGCCCGCCGAGGACCAAGCTCCGAC 111
 Qy 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 Db 112 CAGCGCCCTTCTCTGATAGGGGTGAGCGCGGACCTGCCAGCGGAAGTGCACCTGTGT 171
 Qy 41 GlyLysIleMetGlyLeuLeuGlyGlnAsnGlyValGluGlnArgGlnArgLysVal 60
 Db 172 GAGAAATCATGGAGTTCCTGGGACAGAACAGGTGGAACAGCGGACGAGAGTGTGTC 231
 Qy 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaValAlaLeu 80
 Db 232 ATCTTAGCCAGGACAGGCTTCTAAGAGTCTGACGCGAGAGAGCAAGGCGCAAGGCTTG 291
 Qy 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
 Db 292 AAGGACAGTACATTTTGAACATCCAGATGCTTGTATATATTTGATGACAGAGACT 351
 Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 Db 352 CTGAAGAACTCGTGGAGGCAAAACGGTGGAGGTCCGACCTATGATTTTGTGACACAC 411
 Qy 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 140
 Db 412 TCAAGGTTACAGAGACCAAGGCTCTACCCGCGAGCGTGTGTTGTTGAGGGCATC 471
 Qy 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 Db 472 TTGGGTGTTCAAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCGTGAGACACC 531
 Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 Db 532 GACTCCGAGTCAGGCTGTCTGAAAGATTCTCCGGAGCGTGGCGAGGAGGACCTG 591
 Qy 181 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 200
 Db 592 GAGCAATTCGACGAGATCACCACTTCGTGAACCCGCGCTTCAGAGAGTTCCTCCTG 651
 Qy 201 ProThrLys-LysTyrAlaAspValIleIleProArgGlyValAspAsnMet-ValAlaI 220
 Db 652 CCGACAAAGAGATATGCGATGTGATCATCCACAGAGATGGACAAATATGGGTTGCCA 711
 Qy 220 LeuAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHis- 239
 Db 712 TCAACTGATCGTGGACGACATCCAGACATCTGTAATGTGACATCTGCAATATGGCAC 771
 Qy 240 ArgGlyGlySerAsn-GlyArgSerTyrLysArgThrPheSerGluProGlyAspHis-SP 259
 Db 772 CGAGGAGGTCATGCGGCGGAGCTCAAGCGGACCTTTCGAGCCAGAGACCA--CC 828
 Qy 259 ogLysMetLeuThrSerGlyLysArgSerHis 269
 Db 829 TGGGATGCTGACCTTCGGAACGATCAAT 859
 RESULT 9
 LOCUS B1258532 848 bp mRNA linear EST 17-JUL-2001

DEFINITION	602972340F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112032 5', mRNA sequence.					
ACCESSION	B1258532					
VERSION	B1258532.1 GI:14814971					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	1 (bases 1 to 848)					
TITLE	NIH-MGC http://mgc.nci.nih.gov/.					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapds-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LHAM1271 row: 0 column: 09 High quality sequence stop: 839.					
FEATURES	Location/Qualifiers					
source	1..848					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="IMAGE:5112032"					
	/tissue_type="cervical carcinoma cell line"					
	/lab_host="DH10B"					
	/clone_id="NIH_MGC_12"					
	/note="Organ: cervix; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."					
BASE COUNT	196 a 219 c 271 g 162 t					
ORIGIN						
Alignment Scores:						
Pred. No.:	3,52e-145					
Score:	1300.50					
Percent Similarity:	94.29%					
Best Local Similarity:	93.93%					
Query Match:	89.69%					
DB:	12 Gaps: 1					
US-09-896-522-2 (1-277) x B1258532 (1-848)						
Oy	1 MetAlaSerAlaGlyGluGluProCysGluSerProAlaProGluAlaAspArgProHis 20					
Dd	6 ATGGCTTCGGCGGGAGGCCGAAGACTCGAAGACGGCGCGCGAGCCGACCTCCGAC 65					
Oy	21 GluArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40					
Dd	66 CAGGGGCCCTTCCTGTATAGG-----TCGACCCTGTGT 98					
Oy	41 GluValIleMetGluLeuLeuGlyGlnaAnglValaGluGlnArgGlnArgValVal 60					
Dd	99 GAGAAAGTCAATGAGATTGCTGGGACGAACGAGGGGAAACAGCGGACCGGAAGTGCTC 158					
Oy	61 IleLeuSerGlnaAspArgPheTyrlLysValaLeuThrAlaGluGlnLysAlaLysAlaLeu 80					
Dd	159 ATCTGTAGCCAGGACAGGTTCTTAACAAGTCTCGAAGCGGAGAGCAGAAGCCCAAGCCTTG 218					
Oy	81 LysGlyGlnTyraenPheAspHisPProAspAlaPheAspAsnAPLeuMechisArgThr 100					
Dd	219 AAAGGACAGTCAATTTTGGCCATCCAGATGCCCTTGAATATGATTGATGACAGGACT 278					
Oy	101 LeuLysasnIleValaGluGlyLysThrValaGluValProThrrTyraenPheValThHis 120					
Dd	279 CTGAAGAACATCGTGAAGGGCAAACAGGTGAGAGGTGCCGACCTTGATTTGTTGACACAC 338					

QY	121	SerArgLeuProGluIuThrThValValTyrProAlaAspValValLeuPheGluGlyIle	140
Db	339	TCAGAGTTACAGAGACCAACCGGTGTCTACCTCGGACGGTGTCTGTGTGAGGCGATC	398
QY	141	LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr	160
Db	399	TTGGTGTTCTACAGCAGAGAGATCCGGGACATGTTCCACTCGGCCCTTCTGTGTGGACACC	458
QY	161	AspSerAspValAlaGluSerArgArgValLeuArgAspValAlaGlyGlyValArgAspLeu	180
Db	459	GACTCCGACGTCAGGCTGTCTCGAAGATTCTCCGGGACGTGCCCGCAGGAGGAGCCTG	518
QY	181	GluGlnIleLeuThrGlnTyrThrThrPheValIysProAlaPheGluGluPheCysLeu	200
Db	519	GAGCAGATTCTGACGACCATACACCACTTCGTGGAAGCCGGGCTTTCAGAGAGTTTGTGCTTG	578
QY	201	ProThrIysIysTyrAlaAspValIleIleProArgGlyValaAspAspMetValAlaIle	220
Db	579	CCGACAAAGAAGTATGCCGATGTGTATCATCCACAGAGAGTGGACAAATGTGTGGCATC	638
QY	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsn-GlyAspIleCysIysTyrPheAsp	240
Db	639	AACCTGATCGCGACGACATCTCAGAGACATCTGGAATTGTGACATCTTGCAATGTGCACCG	698
QY	240	GGIAGIYSerAsnGlyArg--SerTyrIysArgThrPheSerGluProGlyAspHisPro	259
Db	659	AGGAGGGTTCATGGGCGGAGGCTTTCACAGCGGACCTTTTGTGACCAAGGAGACCACTT	758
QY	260	GlyMetLeuThrSerGlyIysArg-SerHisLeuGlu-SerSerSerArg	275
Db	759	GGGATGTGACCTCTGGCAACGGGTGCATTGTGAAGTCCAGACGACGAGA	808
RESULT 10			
EM917506			
LOCUS	BM917506	1044 bp	mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT 6606568 NIH_MGC_106	Homo sapiens cDNA clone IMAGE:5484033	
LOCUS	BM917506		5', mRNA sequence.
ACCESSION	BM917506.1	GI:19367885	
VERSION			EST.
KEYWORDS			Homo sapiens (human)
SOURCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE			1 (bases 1 to 1044)
AUTHORS			NIH-MGC http://mgc.nci.nih.gov/ .
TITLE			National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL			Unpublished
COMMENT			Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Dr. Daniel McVicar, DBS/NCI cDNA Library Preparation: Robln Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LNCM2012 row: k column: 10 High quality sequence stop: 624.
FEATURES			Location/Qualifiers
source			1..1044
			/organism="Homo sapiens"
			/mol_type="mRNA"
			/db_xref="taxon:9606"
			/clone="IMAGE:5484033"
			/tissue_type="natural killer cells, cell line"
			/lab_host="DH10B (phage-resistant)"
			/clone_lib="NIH MGC 106"
			/note="Organ: blood; Vector: pOTB7, Site_1: XhoI, Site_2: EcoRI, cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGACGGG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT 224 a 306 c 318 g 196 t

Alignment Scores:

Pred. No.: 2,42e-142 Length: 1044
Score: 1278.00 Matches: 254
Percent Similarity: 95.56% Conservative: 4
Best Local Similarity: 94.07% Mismatches: 8
Query Match: 88.14% Indels: 4
Gaps: 2

US-09-896-522-2 (1-277) x BM917506 (1-1044)

QY 5 G1VGIYGIUASPCYSGIUSERPROIALPROGLIUALAASPARPROHISGINARPROPH 24
DB 1 GGAGGCGAAGACTCGAGAGCCCGCCCGAGGCCACCGTCCGACACCGCCCTTC 60
QY 25 Leuileg1yvalserg1yglthralaserg1yysSerThrValCysGIuylValMet 44
DB 61 CTGATAGGGGTGAGCGCGGCACTGCGAGCGGAACTGACCGTGTGAGAAAGATCATG 120
QY 45 GIuileueng1yGINaeng1yVALGIUINARGLINARGLYVALIleuSerGln 64
DB 121 GAGTGTGTGAGACGAACGAGGTGAAACAGCGGAGGAGGAGGATCTCGAGCCAG 180
QY 65 ASPARGPHEThrylYsVALleuThralaG1uINySAlaYsAlaLeuYsG1yGlnTyr 84
DB 181 GACAGGTTCTACAAAGTCTTGACGCGACGACGAAGGCCCAAGGCTTGAAAGACACTAC 240
QY 85 AsnPhesAspH1sPROAspAlaPheAspAsnAspLeuMetH1sArgThrLeuYsAsnIle 104
DB 241 AATTTGACCATCGAGATGCTTTGATTAATGATTTGATGACAGACTCTGAAGAAACATC 300
QY 105 VALGIUIG1ySvThrValGIUVALProThrThraspPheValThrh1sSerArgLeuPro 124
DB 301 GTGGAGGGCAAAACGGTGAAGGTGCGACCTATGATTTTGACACACTCAAGGTTACCA 360
QY 125 GIuThrThralValTyrProAlaAspValValleuPheGluG1yIleuValPheTyr 144
DB 361 GAGACCAAGGTGCTGCTGCTGCGGAGCTGGTCTGTTTGAAGGCACTTGCTGTTTAC 420
QY 145 SerGlnGluIleArgAspMetPheH1sLeuArgLeuPheValAspThraspSerAspVal 164
DB 421 AGCCAGAGATCCGGGACATGTTCCACTGCGCTTTCGTGACACCGACTCGAGCTC 480
QY 165 ArgLeuSerArgArgValIleuArgAspValArgArgG1yArgAspLeuGluGlnIleu 184
DB 481 AGGCTGTCTGAAGAGTTCTCCGGGAGCTGCGCGGAGGAGGAGCTCGAGCAGATTCTG 540
QY 185 ThrGlnTyrThrThraspValYsProAlaPheGluGluPheCysLeuProThrThrasp 204
DB 541 ACCGACGACACCACTTCTGCGAAGCCGCTTCGAGAGATTTCTGCTCGCCGACAAAGAG 600
QY 205 TyrAlaAspValIleIleProArgG1yValAspAsnMetValAlaIleAsnLeuIleVal 224
DB 601 TATGCCGATGTGATCATCCACGAGGAGTGAACAAATGTTGTCATCAACTGATGCTG 660
QY 225 G1H1sIleG1nAspIleleuAsnG1yAspIleCysIySTPH1sArgG1yG1ySerAsn 244
DB 661 CAGACACTTCAGAGCACTTCGAATGAGATCGCAATTCGAACCAAGAGGATCCCAT 720
QY 245 G1yArg-SerTyrIySArg--ThraspSerGluProG1yAspH1sPro--GlyMetIle 262
DB 721 GGGGGGAGACTACACAGCGGAGACTTTTCTGAGCCCAAGGGAAACCCCTCGGGAGTCT 780
QY 262 u-ThrSerG1yLySArgSerH1sLeu 270
DB 781 GAACCTCTGGGCAAAACGTCCTCACTT 806

RESULT 11
BQ719741 956 bp mRNA linear EST 16-JUL-2002
LOCUS BQ719741
DEFINITION AGENCOUNT_8229663_lupeki_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6184625 5', mRNA sequence.

ACCESSION BQ719741
VERSION BQ719741.1 GI:21858638
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 956)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished
AUTHORS Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

JOURNAL Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: LLAM13574 row: b column: 18
High quality sequence stop: 552.

Location/Qualifiers

FEATURES

source

1..956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6184625"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupeki_dorsal root ganglion"
/note="Vector: PCWt-SPORT (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCCG-3' and
5'-GACTAGTGTGATGCGAGCGGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT 202 a 272 c 296 g 186 t

ORIGIN

Alignment Scores:

Pred. No.: 2.79e-142 Length: 956
Score: 1277.00 Matches: 254
Percent Similarity: 95.19% Conservative: 3
Best Local Similarity: 94.07% Mismatches: 10
Query Match: 88.07% Indels: 3
Gaps: 0

US-09-896-522-2 (1-277) x BQ719741 (1-956)

QY 1 MetAlSerAlaG1yGIUASPCYSGIUSERPROIALPROGLIUALAASPARPROHIS 20
DB 7 ATGGCTTCGGCGGAGGCGAAGCTGCGAGAGCCCGCCCGAGGCCACCGTCCGCAC 66
QY 21 G1NARXPROPHLeuIleG1yValserG1yG1yThralaserG1yysSerThrValCys 40
DB 67 CAGCGGCGCTTCTGTAAGGGGTGAGCGCGGCACTGCGAGGAAAGTGCAGCGTGTGT 126
QY 41 GIuYsIleMetGIuileueng1yGINaeng1yVALGIUINARGLINARGLYVALI 60
DB 127 GAGAGATCATGAGATTGCTGGGACAGAAACAGGTGAAACAGCGGAGAGGTGTC 186

Qy 61 ILeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLeu 80
Db 187 ATCCGAGCCAGGAGAGGTTCTACAGAGTCCGAGCGGAGAGGAGCCAGGCCCTTG 246
Qy 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
Db 247 AAAGACAGGTCAATTTTGACCATCCAGATGCTTTGTAATGATTATCATCAGCGAGCT 306
Qy 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
Db 307 CTGAGAACATCTGAGAGGCGAAGCGGTGGAGGTGCCAGCCTATGATTGTGTGACACAC 366
Qy 121 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyTle 140
Db 367 TCAAGTTACCAAGAGACACAGGTGGTCTACCTCGGAGCGTGGTTCTGTTTGAGGGCATC 426
Qy 141 LeuValPheTyrSerGlnGlnLysLeuArgAspMetPheHisLeuArgLeuPheValAspThr 160
Db 427 TTGGTGTCTACAGCGAGAGATCCGGACATGTTCCACCTGCCCTTGTGTGACACCC 486
Qy 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
Db 487 GACTCCGAGCTGAGGCTCTCTGAGAGATCTCCGGAGCGTGGCGCGAGGAGGAGCCTG 546
Qy 181 GluGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGluGlnLysPheCysLeu 200
Db 547 GAGCAGATTCGAGCGACATACACACCTTCGTGAAGCGGGCCTTGAGAGATTCGCTG 606
Qy 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetVal-AlaI 220
Db 607 CCGCAAGAGATGAGCGATGTGATCCCGCAGAGATGACAAATATGTTTGCCAT 666
Qy 220 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArg 240
Db 667 CAACCTGATCTGACGACATCCAGAGCTTCTTATGGGAACTCTGCAATATGGCACCC 726
Qy 240 gGlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGluPro-GlyAspHisPro 259
Db 727 AGGAAGGTCCAATGGGCGGACCTCCAAAGGAGACCTTTCTTAACCAAGGAGAACCCCT 786
Qy 260 GlyMetLeuThrSerGlyLysArg 267
Db 787 GGGATGTTGACCTTGCCCTAAG 810

RESULT 12
BUS11803 943 bp mRNA linear EST 12-SEP-2002
LOCUS AGNCOURT_10116119 NIH_MGC_134 Mus musculus cDNA clone
DEFINITION IMAGE:6506624 5', mRNA sequence.
ACCESSION BUS11803
VERSION BUS11803.1 GI:22818036
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 943)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-romail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM4069 row: g column: 09
High quality sequence stop: 683.
Location/Qualifiers
1..943

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6506624"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 134"
/note="Vector: pCMV-SPORT6.1, cdb; Site 1: EcoRV; Site 2:
NotI; Cloned unidirectionally. Primer: Oligo dT. Average
insert size 1.7 kb. Constructed by Resgen, Invitrogen
Corp. Note: This is a NIH MGC Library."
BASE COUNT 221 a 263 c 277 g 181 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 6 25e-142 Length: 943
Score: 1274.00 Matches: 247
Percent Similarity: 94.89% Conservative: 13
Best Local Similarity: 90.15% Mismatches: 12
Query Match: 87.86% Indels: 2
DB: 13 Gaps: 0
US-09-896-522-2 (1-277) x BUS11803 (1-943)
Qy 6 GlyLysAspCysGluSerProAlaProGluAlaAspArgProHisGlnArgProPheLeu 25
Db 3 GGGCGCGCTCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 62
Qy 26 ILeGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLysIleMetGlu 45
Db 63 ATGGGCTGACGGCG 122
Qy 46 LeuLeuGlyValAsnGluValGluGlnArgGlnArgLysValValIleLeuSerGlnAsp 65
Db 123 CTGTGGGACAGAACGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
Qy 66 ArgPheTyrLysValLeuThrAlaGluGlnLysAlaLeuLysGlyGlnTyrAsn 85
Db 183 TGCTTCAAGAGTTCTGACGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
Qy 86 PheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleVal 105
Db 243 TTTGACCAACCAAGATGCTTTGTAATGATCATGACAAAGACCTTGAAAAACATTTGT 302
Qy 106 GluGlyLysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGlu 125
Db 303 GAAGGCAAACTGTGAGGTCCCTACCTATGATTTGTGACCACTCAAGGTTTACAGAG 362
Qy 126 ThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSer 145
Db 363 ACCACTGTGTCTACCAAGCTGATGTGTTCTGTTGAGGGCATCTTGGTATTCTACACC 422
Qy 146 GlnGluLysArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArg 165
Db 423 CAGAGATCCGGAGACATGTTCCACCTGCCCTTTGTGTGACACACAGACTCTGATGTTAGG 482
Qy 166 LeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGlnIleLeuThr 185
Db 483 CTGCTCGAAGATTTCTCCGGATGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
Qy 186 GlnTyrThrThrPheValLysProAlaPheGluGlnLysPheCysLeuProThrLysLysTyr 205
Db 543 CAGTACACCGCCCTTTGTGAAGACAGGCTTTGAGGATTTCTGCTGCGCATAGAGAGTAC 602
Qy 206 AlaAspValIleIleProArgGlyValAspAsnMetValAlaIleAsnLeuIleValGln 225
Db 603 GCTGAGGTGATCATCCCTCGAGAGATGTAATATGTGGCCATCAACCTGATCGTGCAA 662
Qy 226 HisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArgGlyGlySerAsnGly 245
Db 663 CACATCCAGGACATCTTCAAGGAGCTGTGCAAGGAGAGAGAGAGAGAGAGAGAGAGAG 722
Qy 246 ArgSerTyrLysArgThrPheSerGluProGlyAspHisPro-GlyMetLeuThrSerG 265

http://image.llnl.gov
 plate: LHAM1353 row: n column: 22
 High quality sequence start: 38
 High quality sequence stop: 941.
 Location/Qualifiers

FEATURES

source

```

1..961
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 1T"
/db_xref="taxon:10090"
/clone="IMAGE:5143509"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NCI CGAP Lu33"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'.
TGTTACCAATCTGAAGTGGAGCGCGCCCTCTTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
```

BASE COUNT 229 a 252 c 289 g 191 t

ORIGIN

Alignment Scores:

Score: 3.08e-140 Length: 961
 Percent Similarity: 1260.00 Matches: 248
 Best Local Similarity: 92.45% Conservative: 9
 Query Match: 86.90% Mismatches: 20
 Indels: 1
 Gaps: 0

US-09-896-522-2 (1-277) x B1416102 (1-961)

```

Qy 1 MetlaserlaaglygluaspCyguSerProalaProglualaaspArgProhis 20
Db 63 ATGGCTTCGGGGAGGGGGGGCGCTCGAGAGCCGCCGCCGAGCGATGCCAG 122
Qy 21 GluArgProPheleuileglyvalserglylythrAlaSerglylyserThrValCys 40
Db 123 CCGGGCGCTTCCATCGCGGTGAGCGCGCGCACCCGCTAGTGCACAGTGTGT 182
Qy 41 GluysilemetguileuileuglygluasngluvalgluInarggluInarglyval 60
Db 183 GAGAAAGATCAGGAGCTGCTGGACAGACGAAAGTGGACCCGCCGACGCAAGTGTG 242
Qy 61 lleuSerGlnaSPArgPheTyrlYserValleuthrAlaIngluInlyAlaValaleu 80
Db 243 ATCTGAGCCAGGAGCTGCTTCAACAAGTTTGACGGCGGAGCAAGAGCCAGCTTGG 302
Qy 81 lyselYglInlyrAanPheaspHisProaspAlaPheaspAnaSPleuMetHisArgThr 100
Db 303 AAGGAGACAGTCAATTTTGACACCCAGATCTTTGATATGATCTGATGACAAAGCC 362
Qy 101 leuysanileValIgluIlyleThrValgluValProthryrAspPheValThHis 120
Db 363 CTGAAAACATTTGTAAGGCAAACTGTCCAGGCTCCCTACTATGATTTTGTACCCAC 422
Qy 121 SerArgleuProgluThrThrValValTyrrProAlaaspValIleuPhegluIlyle 140
Db 423 TCAAGTTTACAGAGACCACTGTGTCTACCCAGTATGTGTCTGTTGAGGGCATC 482
Qy 141 leuValPheTyrlYserGlnIlyleArgaspMetPheHisleuArgleuPheValAspThr 160
Db 483 TTGTTATTTCTACACCCAGAGATCCGGGACATGTTCCACCTCGCCCTTTGTGACACA 542
Qy 161 AspSeraspValaArgleuSerArgArgValleuArgaspValaArgagglyArgAspLeu 180
Db 543 GACTCTGATGTTAGGCTGTCTCGAAGAGTTCTCCGGGATGTGCAACAGAGGAGGACCTG 602
```

```

Qy 181 GluGlnIleuThrArgInlyrThrThrPheValIysProAlaPhegluInuPheCysLeu 200
Db 603 GAGAGATCTCTGACTAGTACACCGCCTTTGTGAACACGCTTTAGAGAGTCTGCTG 662
Qy 201 ProThrIylsTyrlAlaaspValIleIleProArgglyValAspAnMetValAlaIle 220
Db 663 CCACACTAAGAAATGACGTGACGTGATCATCTCCGAGAGAGTTATATATGTGGCATC 722
Qy 221 AspleuileValIghnIstIleGlnaspIleleuAspGlyAspIleCysIlySTPHisArg 240
Db 723 AACCTGATCTGGAACATCCAGACATCTTCAACGGGACCTGTGCAAGCGGACCGA 782
Qy 241 GlyIySerAnsglyArg-SerTyrlYsArgThrPheSerGluProglYAspHisProgl 260
Db 783 GGGGGGGCAACGGGGGCAACAAACAGAGACTTCCCGGACCGAGAGATCACCTGGG 842
Qy 260 YmetleuthrSerGlylyleArgSerHisleuGluSerSerSerArgProhis 277
Db 843 GTGTTCCGCACTTGGCAAGCGCTTCAACATGAGTGTAGCAGAGAACCCAT 894
```

RESULT 15

BG390519

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

602416183F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4524308 5', mRNA sequence.
 BG390519 GI:13283967
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 989)
 NIH-MGC http://mgc.nci.nih.gov/
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LHAM10428 row: f column: 21
 High quality sequence stop: 750.
 Location/Qualifiers

FEATURES

source

```

1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4524308"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."
```

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 4.24e-140 Length: 989
 Score: 1259.00 Matches: 258
 Percent Similarity: 94.89% Conservative: 2
 Best Local Similarity: 94.16% Mismatches: 9
 Query Match: 86.83% Indels: 7
 Gaps: 1

US-09-896-522-2 (1-277) x BG390519 (1-989)

```
QY 7 GluAspCysGluSerProAlaProGluAlaAspArgProHisGluAsnArgProPheLeuIle 26
   |||||
Db 2 GAAGACTGGCAGAGACCCCGCGCCGAGAGCCGACCTCCGACCGAGCCGCTTCTGATA 61
QY 27 GlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGluLysIleMetGluLeu 46
   |||||
Db 62 GGGGTGAGCGCGGCACTGCGCAGCGGAGAGTCGACCGTGTGTGAGGAAGATCATGAGATTG 121
QY 47 LeuGlyGlnAsnGluValGluGlnArgGluAsnArgLysValValIleLeuSerGluAspArg 66
   |||||
Db 122 CTGGGACAGAACGAGCGAGAACAGCGGACGCGAAGGTGTCATCTGAGCCAGACAGCAGG 181
QY 67 PheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeuLysGlyGlnTyrAsnPhe 86
   |||||
Db 182 TTCTACAAAGTCTCTGACGCGACAGACAGAAAGCCCAAGCCTTGAAAGACAGATCAATTTT 241
QY 87 AspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeuLysAsnIleValGlu 106
   |||||
Db 242 GACCATCCAGATGCTTTGATTAATGATTGATGACAGAGACTCGAAGAACATGTGGAG 301
QY 107 GlyLysThrValGluValProThrTyrAspPheValThrHisSerArgLeuProGluThr 126
   |||||
Db 302 GGCAGAAACGGTGGAGTGCAGCTATGATTGTGACACATCAAGTTACAGAGACC 361
QY 127 ThrValValTyrProAlaAspValValLeuPheGluGlyIleLeuValPheTyrSerGln 146
   |||||
Db 362 ACGGTGCTTACCTCGCGAGCGTGTCTGTTTGAGGACACTTGTGTCTTACAGCCAG 421
QY 147 GluIleArgAspMetPheHisLeuArgLeuPheValAspThrAspSerAspValArgLeu 166
   |||||
Db 422 GAGATCCGGGACATGTTCCACCTCGCGCTCTGTGTGACACCGACTCCGAGCTGAGGCTG 481
QY 167 SerArgArgValLeuArgAspValArgArgGlyArgAspLeuGluGlnIleLeuThrGln 186
   |||||
Db 482 TCTCGAAGAGTTCCTCCGCGACGTGCGCGAGGAGGAGCCTGGAGCAGATTCTGACCGCAG 541
QY 187 TyrThrThrPheValLysProAlaPheGluGluPheCysLeuProThrLysValTyrAla 206
   |||||
Db 542 TACACCACTTCGTGAAAGCCGAGCTTCGAGAGTTCGCTCGCGACAAAGAGTATGCC 601
QY 207 AspValIleIleProArgGlyVal-AspAsnMetVal-AlaIleAsnLeuIleValGlnH 226
   |||||
Db 602 GATGTGATCATCCCGCAGAGGAGTTCGACAAATATGTTGGCCATCAAGCTGATGTGCAGC 661
QY 226 IsoleuAsnAspIleLeuAsnGlyAspIleCysLys-TrpHisArg-GlyLysSerAsnG 245
   |||||
Db 662 ACATCCAGGACATTCTGAAATGTGACATCTGCAATTGGCAGCAGAAAGATCCAAATGG 721
QY 245 YArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlyMetLeuThrSerG 265
   |||||
Db 722 GCGGAGCTTACAGGCGACTTTTCT-GAGCA-GGAGACCACCCGTGATTTGCTGACCTCTGG 779
QY 265 YLysArgSerHisLeuGluSerSerSerArgPro 276
   |||||
Db 780 C---AACGGGACACATTGGAGCCACGACAGACC 810
```

Search completed: November 25, 2003, 10:07:08
Job time : 2563 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:06:00 ; Search time 39 Seconds
(without alignments)
1127.366 Million cell updates/sec

Title: US-09-896-522-2

Sequence: 277
1 MASAGDEDESPAPADRP.....HPGMLTSGKSHLESSSRH 277

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 15

Total number of hits satisfying chosen parameters: 16

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.GeneSeq.19Jun03:*
1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	277	22	AA199502
2	277	100.0	277	22	AA199502
3	277	100.0	277	22	AA199502
4	277	100.0	277	22	AA199502
5	277	100.0	277	22	AA199502
6	277	100.0	277	22	AA199502
7	277	100.0	277	22	AA199502
8	277	100.0	277	22	AA199502
9	277	100.0	277	22	AA199502

10	25	9.0	261	22	AA199502	Human polypeptide
11	25	9.0	261	22	AA199502	Human transferrin
12	25	9.0	337	21	AA199502	Human prostate can
13	25	9.0	337	21	AA199502	Human ovarian anti
14	15	5.4	125	23	AA199502	Human 57658 protei
15	15	5.4	260	22	AA199502	Drosophila melanog
16	15	5.4	335	22	AA199502	Human polypeptide

ALIGNMENTS

RESULT 1
AA199502
ID AA199502 standard; Protein; 277 AA.
XX
AC AA199502;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2647.
XX
KW Human; noctropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW Leukemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WC-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSER INC.
XX
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX MPI: 2001-442253/47.
XX N-PSDB; AA158658.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX
XX Example 4; SEQ ID NO 2647; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AA158642-AA164213) with noctropic,
XX immunosuppressant and cytoskeletal activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 277 AA;
Query Match 100.0%; Score 277; DB 22; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.6e-256;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGDECSPEADRPQHQPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
Db 1 MASAGDECSPEADRPQHQPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
QY 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLKNIVEKTVVPTDYVTH 120
Db 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLKNIVEKTVVPTDYVTH 120
QY 121 SRLPETTVVYPADVLPFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
Db 121 SRLPETTVVYPADVLPFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVII PRGVDMNVAIINLVQHIQDILNGDICKMR 240
Db 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVII PRGVDMNVAIINLVQHIQDILNGDICKMR 240
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
Db 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
RESULT 2
AAB93941
ID AAB93941 standard; Protein; 277 AA.
XX
AC AAB93941;
XX
DT 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:13952.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto U;
XX
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX
XX Claim 8; SEQ ID 13952; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 277 AA;
Query Match 100.0%; Score 277; DB 22; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.6e-256;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAGDECSPEADRPQHQPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
Db 1 MASAGDECSPEADRPQHQPFLIGVSGGTASGKSTVCEKIMELLGONEVEQORQKV 60
QY 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLKNIVEKTVVPTDYVTH 120
Db 61 ILSODRFYKVLTAQKAKALKGQYNFHPDAPFNDLMHRTLKNIVEKTVVPTDYVTH 120
QY 121 SRLPETTVVYPADVLPFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
Db 121 SRLPETTVVYPADVLPFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
QY 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVII PRGVDMNVAIINLVQHIQDILNGDICKMR 240
Db 181 EQLITQYTTFFVKPAFEFECFLPTKKYADVII PRGVDMNVAIINLVQHIQDILNGDICKMR 240
QY 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
Db 241 GGSNGRSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
RESULT 3
AAB16592
ID AAB16592 standard; Protein; 277 AA.
XX
AC AAB16592;
XX
DT 18-APR-2002 (first entry)
DE Human 57658 protein.
XX
XX Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
XX 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
XX rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
XX sarcoma; myocardial infarction; hypertension; atherosclerosis;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease;
XX gene therapy; chromosome mapping; tissue typing; dermatological;
XX cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
XX anticonvulsant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 5..10
XX FT /label= N-myristoylation_site
XX FT Domain 25..151

FT		/note= "kinase uridine transferase ATP-binding phosphoribulokinase monophosphokinas precursor PRK cycle phosphopentokinase domain"
FT		25..231
FT	Domain	/note= "phosphoribulokinase domain; Protein kinase family domain"
FT		27..32
FT	Modified-site	/label= N-myristoylation_site
FT		30..37
FT	Binding-site	/label= ATP/GTP-binding_site
FT		34..36
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		38..41
FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
FT		100..102
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		108..115
FT	Modified-site	/label= Tyrosine_kinase_phosphorylation_site
FT		154..277
FT	Domain	/note= "kinase uridine monophosphokinas transferase ATP-binding kinase-like ribonucleoside pyrimidine FIS cDNA domain"
FT		160..163
FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
FT		167..169
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		202..204
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		241..246
FT	Modified-site	/label= N-myristoylation_site
FT		247..249
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		251..254
FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
FT		260..265
FT	Modified-site	/label= N-myristoylation_site
FT		264..266
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
FT		264..267
FT	Modified-site	/label= Amidation_site
FT		268..271
FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
FT		273..275
FT	Modified-site	/label= Protein_kinase-C_phosphorylation_site
PX		WO200202761-A2.
PN		
XX		
PD		10-JAN-2002.
XX		
PF		28-JUN-2001; 2001WO-US21063.
PR		30-JUN-2000; 2000US-216503P.
PA	(MILL-) MILLENNIUM PHARM INC.	
PI	Gluckemann MA;	
DR	WPI: 2002-140091/18.	
XX	N-PSDB; AAD27186.	
PT		
PT	New isolated human uridine kinase family polypeptide 57658, useful for treating hematopoietic neoplastic disorders and disorders of neurons, heart and blood vessels _	
PS	Claim 4; Fig 1a; 103pp; English.	
CC	The patent discloses human uridine kinase-like polypeptides, designated 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins are useful for developing diagnostic and therapeutic agents for 57658-mediated or related disorders such as haematopoietic neoplastic disorders (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis, diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets, osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels	

Query Match	100.0%	Score 277	DB 23	Length 277
Best Local Similarity	100.0%	Pred. 2.6e-256		
Matches 277	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MASAGGECSEPPAEADRPHPORPPLIGVSGGTASGKSTVEKIMELLGONEVEORQRKYV	60	
DB	1	MASAGGECSEPPAEADRPHPORPPLIGVSGGTASGKSTVEKIMELLGONEVEORQRKYV	60	
QY	61	ILSDRPFVKVLTAQKAKALKGQVNFHPDAFNDLMMHRTLKNTVEGKTVEVPYTPDFVTH	120	
DB	61	ILSQDRFVKVLTAQKAKALKGQVNFHPDAFNDLMMHRTLKNTVEGKTVEVPYTPDFVTH	120	
QY	121	SRLEPPTVVYVADAVLFEGLIVFVFSOEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL	180	
DB	121	SRLEPPTVVYVADAVLFEGLIVFVFSOEIRDMFHLRLFPVDTSDVRLSRVLRDVRGRDL	180	
QY	181	EQILTYTTFVKPAPEEFCELPTRYKXADYIIPRGVDNMVAIQLVQHIDILINDGICCKWR	240	
DB	181	EQILTYTTFVKPAPEEFCELPTRYKXADYIIPRGVDNMVAIQLVQHIDILINDGICCKWR	240	
QY	241	GGSNGRSYKRTFSPGDPHGLTSGKSHLESSSRPH	277	
DB	241	GGSNGRSYKRTFSPGDPHGLTSGKSHLESSSRPH	277	
RESULT 4				
AA041288	AA041288 standard; Protein; 296 AA.			
XX	AC	AA041288;		
XX	DT	22-OCT-2001 (first entry)		
DE	XX	Human polypeptide SEQ ID NO 6219.		
XX	XX	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	KW	chemokineic; thrombolytic; drug screening; arthritis; inflammation;		
XX	OS	Homo sapiens.		
XX	PN	WO200153312-A1.		
XX	PD	26-JUL-2001.		
XX	PF	26-DEC-2000 in 2000WO-US34263.		
XX	PR	21-JAN-2000; 2000US-0488725.		
PR	PR	25-APR-2000; 2000US-0552317.		
PR	PR	09-JUL-2000; 2000US-0598042.		
PR	PR	19-JUL-2000; 2000US-0620312.		
PR	PR	03-AUG-2000; 2000US-0653450.		
PR	PR	14-SEP-2000; 2000US-0662191.		
PR	PR	19-OCT-2000; 2000US-0693036.		
PR	PR	29-NOV-2000; 2000US-0727344.		
PA	(HYSE-)	HYSEQ INC.		

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang Z, Wang Z, Weinman T, Xu C, Xue A, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB; AA160444.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 2; SEQ ID NO 6219; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA038642-AA042213) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilization of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX
 SQ Sequence 296 AA;
 Query Match 100.0%; Score 277; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 2,8e-256;
 Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGDESPAEADRPORFLIGVSGTASGKSTVCCKIMELLGONEVEQRORRV 60
 Db 20 MASAGGDESPAEADRPORFLIGVSGTASGKSTVCCKIMELLGONEVEQRORRV 79
 QY 61 ILSQDRFYKVLTAQAKALKQYNFDPDAFNDLMHRTLKNIVEGTVFPTVDFVTH 120
 Db 80 ILSQDRFYKVLTAQAKALKQYNFDPDAFNDLMHRTLKNIVEGTVFPTVDFVTH 139
 QY 121 SRLPFTTVVPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 Db 140 SRLPFTTVVPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 199
 QY 181 EQILTOYTTFFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
 Db 200 EQILTOYTTFFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 259
 QY 241 GGSNGRSYKRTFSEPGDHGMLTSGKRSHTLESSSRPH 277
 Db 260 GGSNGRSYKRTFSEPGDHGMLTSGKRSHTLESSSRPH 296
 Db
 RESULT 5
 AA014412
 ID AA014412 standard; Protein; 260 AA.
 XX
 AC AA014412;
 XX
 DT 02-MAY-2002 (first entry)
 XX
 DE Protein of a human uridine kinase (UDK).
 XX
 KW Human: uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen.
 XX
 OS Homo sapiens.

XX
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI: 2001-626259/72.
 DR N-PSDB; AAK98735.
 XX
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance -
 XX
 PS Claim 3; Page 29-30; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilizing uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example, detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome
 CC localisation studies and tissue expression studies. The polypeptides of
 CC the invention or fragments thereof may be used as immunogens to produce
 CC antibodies. These antibodies may be employed to isolate or identify
 CC clones expressing the polypeptide. The polypeptides and polynucleotides
 CC of the invention can be used as a vaccine or in gene therapy to treat
 CC diseases such as human ovarian cancer, human colon carcinomas, and
 CC immunological disorders. This sequence represents the protein of a human
 CC uridine kinase of the invention.
 CC
 XX
 SQ Sequence 260 AA;
 Query Match 93.9%; Score 260; DB 22; Length 260;
 Best Local Similarity 100.0%; Pred. No. 4.4e-240;
 Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASAGGDESPAEADRPORFLIGVSGTASGKSTVCCKIMELLGONEVEQRORRV 60
 Db 1 MASAGGDESPAEADRPORFLIGVSGTASGKSTVCCKIMELLGONEVEQRORRV 60
 QY 61 ILSQDRFYKVLTAQAKALKQYNFDPDAFNDLMHRTLKNIVEGTVFPTVDFVTH 120
 Db 61 ILSQDRFYKVLTAQAKALKQYNFDPDAFNDLMHRTLKNIVEGTVFPTVDFVTH 120
 QY 121 SRLPFTTVVPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 Db 121 SRLPFTTVVPADVLFEGILVFYSQEI RDMFHLRLFVDTSDVRLSRVLRDVRGRDL 180
 QY 181 EQILTOYTTFFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
 Db 181 EQILTOYTTFFVKPAFEFECPLPTKKYADVILIPGVDMVAIINLVQHIQDILNGDICKMHR 240
 QY 241 GGSNGRSYKRTFSEPGDHG 260
 Db 241 GGSNGRSYKRTFSEPGDHG 260
 Db
 RESULT 6
 AAG64506

```

ID  AAG64506 standard; Protein; 276 AA.
XX
AC  AAG64506;
XX
DT  02-OCT-2001 (first entry)
XX
DE  Human uridine kinase.
XX
KW  Human; uridine kinase; UK.
XX
OS  Homo sapiens.
XX
PN  CN1287172-A.
XX
PD  14-MAR-2001.
XX
PF  07-SEP-1999; 99CN-0118818.
XX
PR  07-SEP-1999; 99CN-0118818.
XX
PA  (UYFU-) UNIV FUDAN.
XX
PI  Yu L, Zhao Y, Zhang H;
XX
DR  WPI; 2001-409529/44.
DR  N-PSDB; AAH75355.
XX
PT  Human uridine kinase and its coding sequence, preparation and
XX  application -
XX
PS  Claim 2; Page 15-16(Disclosure); 20pp; Chinese.
XX
CC  The invention relates to human uridine kinase (UK).
XX
SQ  Sequence 276 AA;

Query Match      68.2%; Score 189; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.3e-172;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  58 KVVILSQDRFYKVLTAEQKAKALKGQYNFDPDAFNDMLMRTLKNIVEGKTYEVPYYDF 117
    |||||
DB  57 KVVILSQDRFYKVLTAEQKAKALKGQYNFDPDAFNDMLMRTLKNIVEGKTYEVPYYDF 116

QY  118 VTHSRLEPETYVYPADVILFEGILVFYSQETRDMFHRLFYDITDSVRLSRVLRDVRG 177
    |||||
DB  117 VTHSRLEPETYVYPADVILFEGILVFYSQETRDMFHRLFYDITDSVRLSRVLRDVRG 176

QY  178 RDLFQILTOYTTFPKPAFEFCLPTKKYADYIIPRGVDNMVAIYLIVQHIDILNGDICK 237
    |||||
DB  177 RDLFQILTOYTTFPKPAFEFCLPTKKYADYIIPRGVDNMVAIYLIVQHIDILNGDICK 236

QY  238 WHRGSGNGR 246
    |||||
DB  237 WHRGSGNGR 245

RESULT 7
ABB89353
ID  ABB89353 standard; Protein; 190 AA.
XX
AC  ABB89353;
XX
DT  24-MAY-2002 (first entry)
XX
DE  Human polypeptide SEQ ID NO 1729.
XX
KW  Cytosolic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX  antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX  vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX  cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX  neurological disease; infection; human; secreted protein.
XX

```

```

OS  Homo sapiens.
XX
PN  NO200190304-A2.
XX
PD  29-NOV-2001.
XX
PF  18-MAY-2001; 2001WO-US16450.
XX
PR  19-MAY-2000; 2000US-205515P.
XX
PA  (HUMA-) HUMAN GENOME SCI INC.
XX
PI  Birse CE, Rosen CA;
XX
DR  WPI; 2002-122018/16.
DR  N-PSDB; ABL89762.
XX
PT  Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT  prevention of neural, immune system, muscular, reproductive,
PT  gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT  disorders -
XX
PS  Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
XX
CC  The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC  (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC  medical conditions e.g. by protein or gene therapy. The genes are
CC  isolated from a range of human tissues disclosed in the specification.
CC  The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC  in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC  and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC  marrow, breast, gastrointestinal tract, liver, lung or urogenital;
CC  (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC  haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC  disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC  colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC  (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC  epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC  and parasitic infections.
CC  Note: The sequence data for this patent did not form part of the
CC  printed specification, but was obtained in electronic format directly
CC  from WIPO at fcp.wipo.int/pub/published_poc_sequences.
XX
SQ  Sequence 190 AA;

Query Match      49.1%; Score 136; DB 23; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.1e-121;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34 SGKSTYCEKIMELLGONEVEORQRKVVILSQDRFYKVLTAEQKAKALKGQYNFDPDAF 93
    |||||
DB  34 SGKSTYCEKIMELLGONEVEORQRKVVILSQDRFYKVLTAEQKAKALKGQYNFDPDAF 93

QY  94 NDLMHRITLKNIVEGKTYEVPYDFTVTHSRLEPETYVYPADVILFEGILVFYSQETRDMFH 153
    |||||
DB  94 NDLMHRITLKNIVEGKTYEVPYDFTVTHSRLEPETYVYPADVILFEGILVFYSQETRDMFH 153

QY  154 LRLFVDTSDVRLSRR 169
    |||||
DB  154 LRLFVDTSDVRLSRR 169

RESULT 8
ABG70328
ID  ABG70328 standard; Protein; 120 AA.
XX
AC  ABG70328;
XX
DT  21-OCT-2002 (first entry)
XX
DE  Human MBDT protein Incyte ID No. LI:235557.12.orf2:2001JAN12.
XX
KW  Human; molecule for disease detection and treatment; MBDT; cancer;
XX

```

KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cystostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiaschmatic.
XX Homo sapiens.
XX WO200255738-A2.
XX
XX 18-JUL-2002.
XX
XX 09-JAN-2002; 2002WO-US01008.
XX
XX 12-JAN-2001; 2001US-261622P.
XX 16-JAN-2001; 2001US-261865P.
XX 17-JAN-2001; 2001US-262208P.
XX 17-JAN-2001; 2001US-262209P.
XX 17-JAN-2001; 2001US-262326P.
XX 19-JAN-2001; 2001US-263063P.
XX 19-JAN-2001; 2001US-263065P.
XX 19-JAN-2001; 2001US-263329P.
XX
XX (INCYTE GENOMICS INC.
XX Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AJ,
XX Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ,
XX Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA,
XX
XX WPI; 2002-590679/63.
XX N-PSDB; ABS51801.
XX
XX New disease detection and treatment molecule (MDPT) polynucleotides and
XX polypeptides, useful in diagnosing, studying, preventing or treating
XX diseases associated with MDR expression, e.g. autoimmune or
XX inflammatory disorders -
XX
XX Claim 27; Page 120; 123pp; English.
XX
XX The present invention relates to the isolation of novel human
XX molecules for disease detection and treatment (MDPT), and the
XX polynucleotide sequences (mdpt) encoding them. The MDR polypeptides
XX may be used to screen for molecules that bind to, or are bound by the
XX encoded polypeptides, and to develop a transcript image of a tissue or
XX cell type. Probes comprising at least 20 nucleotides of the mdpt
XX polynucleotide may be used to assess the toxicity of a test compound.
XX The MDR polypeptides and mdpt polynucleotides are useful in the
XX diagnosis, study, prevention and treatment of diseases associated with
XX the expression of molecules for disease detection and treatment. Such
XX disorders include cell proliferative disorders (e.g. arteriosclerosis,
XX cirrhosis, or cancers), and autoimmune/inflammatory disorders
XX (e.g. asthma, Crohn's disease, or multiple sclerosis). The mdpt
XX polynucleotides may also be used as molecule markers, in microarrays,
XX and in somatic or germline gene therapy. ABS70306-ABS70341 represent
XX the MDR proteins of the invention.
XX
XX Sequence 120 AA;
XX
XX Query Match 40.4%; Score 112; DB 23; Length 120;
XX Best Local Similarity 100.0%; Pred. NO. 6.4e-99;
XX Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 90 DAFNDNDLHRTLNKIVGKTVVPTVDFVTHSRLEPFTVVYPADVLFEGILVFSQSEIR 149
XX DAFNDNDLHRTLNKIVGKTVVPTVDFVTHSRLEPFTVVYPADVLFEGILVFSQSEIR 64
XX 5 DAFNDNDLHRTLNKIVGKTVVPTVDFVTHSRLEPFTVVYPADVLFEGILVFSQSEIR 64
XX
XX 150 DMFHLRLFVDTSDVRLSRVLRDVRGRDLLEQLITQTTFTVKAFEEFCIP 201
XX DMFHLRLFVDTSDVRLSRVLRDVRGRDLLEQLITQTTFTVKAFEEFCIP 116
XX
XX RESULT 9
XX AAO14413
XX ID AAO14413 standard; Protein; 277 AA.

XX
XX AAO14413;
XX
XX 02-MAY-2002 (first entry)
XX
XX Protein relating to a human uridine kinase (UDK) of the invention.
XX
XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
XX probe; chromosome localisation study; tissue expression; gene therapy;
XX antibody; vaccine; human ovarian cancer; immunological disorder;
XX human colon carcinoma; immunogen.
XX
XX Unidentified.
XX
XX WO200172963-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMITIK) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
XX and to identify agonists and antagonists of the polypeptide useful for
XX treating conditions associated with uridine kinase imbalance -
XX
XX Disclosure; Page 23; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
XX polypeptides and polynucleotides and methods for producing such
XX polypeptides by recombinant techniques. Also disclosed in the invention
XX are methods for utilising uridine kinase polypeptides and polynucleotides
XX in diagnostic assays. The polynucleotides and polypeptides of the
XX invention may be used as diagnostic reagents by detecting mutations in an
XX associated gene. An array of oligonucleotide probes comprising the
XX uridine kinase polynucleotide sequence or fragments thereof can be
XX constructed to conduct efficient screening of genetic mutations, for
XX example. Detection of abnormally decreased or increased levels of
XX polypeptide or mRNA expression may also be used for diagnosing or
XX determining susceptibility of a subject to a disease of the invention.
XX The polynucleotide sequences of the invention can be used for chromosome
XX localisation studies and tissue expression studies. The polypeptides of
XX the invention or fragments thereof may be used as immunogens to produce
XX antibodies. These antibodies may be employed to isolate or identify
XX clones expressing the polypeptide. The polypeptides and polynucleotides
XX of the invention can be used as a vaccine or in gene therapy to treat
XX diseases such as human ovarian cancer, human colon carcinomas, and
XX immunological disorders. This sequence represents the protein relating to
XX a human uridine kinase (UDK) of the invention.
XX NOTE: The present sequence is stated as being the same as that shown as
XX SEQ ID NO: 2 in the sequence listing of the specification. However, the
XX sequences differ.
XX
XX Sequence 277 AA;
XX
XX Query Match 16.2%; Score 45; DB 22; Length 277;
XX Best Local Similarity 100.0%; Pred. NO. 1.4e-34;
XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 100 TLKNIVGKTVVPTVDFVTHSRLEPFTVVYPADVLFEGILVY 144
XX TLKNIVGKTVVPTVDFVTHSRLEPFTVVYPADVLFEGILVY 144
XX
XX RESULT 10
XX AAM38694
XX ID AAM38694 standard; Protein; 261 AA.

XX AAM38694;
AC 22-OCT-2001 (first entry)
DT Human polypeptide SEQ ID NO 1839.
XX

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.

XX Homo sapiens.
PN WO200153312-A1.
XX

XX 26-JUL-2001.
XX

XX 26-DEC-2000; 2000MO-US34263.
XX

XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.
XX

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX

DR WPI: 2001-442253/47.
XX N-PSDB; AA157850.
XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX

PS Example 3; SEQ ID NO 1839; 10078bp; English.
XX

XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AA42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 261 AA;
SQ

Query Match 9.0%; Score 25; DB 22; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECPLTKKADVIIRG 213
DB 187 TFWKPAFEFECPLTKKADVIIRG 211

RESULT 11

ID AAB73494 standard; Protein; 261 AA.
XX

AC AAB73494;
XX

DT 31-JUL-2001 (first entry)
XX

XX Human transferase HTFS-1, SEQ ID NO:1.
XX

XX Human transferase; HTFS; agonist; antagonist; cellular signalling;
KM proliferation; cell proliferative disorder; immune disorder;
KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KM gene therapy; drug screening.

XX Homo sapiens.
OS

XX WO200132888-A2.
PN

XX 10-MAY-2001.
PD

XX 02-NOV-2000; 2000MO-US30485.
PF

XX 04-NOV-1999; 99US-0163595.
PR

XX (INCYTE-) INCYTE GENOMICS INC.
PA

XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih LT, Azimzai Y, Lu DM, Baughn MR;
XX

DR WPI: 2001-328796/34.
XX N-PSDB; AAH23801.
XX

XX Human transferase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX

PS Claim 1; Page 103-104; 157bp; English.
XX

XX Sequences AAB73494-AAB73535 represent novel human transferase proteins
XX HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
XX encoding them. The proteins play important roles in the regulation of
XX cellular signalling and proliferation. The HTFS proteins are useful for
XX screening compounds for their effectiveness as agonists or antagonists of
XX transferase activity, or for compounds that specifically bind to an HTFS
XX protein or which modulates the activity of an HTFS protein.
XX Pharmaceutical compositions comprising an HTFS protein, HTFS
XX agonist or antagonist, or genetic construct encoding an HTFS
XX protein are useful for treating a disease or condition associated
XX with decreased or increased expression of functional HTFS disorders
XX which may be treated using such compositions include cell proliferative
XX disorders and immune disorders. For example, diseases which may be
XX treated include atherosclerosis, hepatitis, psoriasis, cancers (including
XX breast, bladder, bone marrow, brain and uterus cancer), inflammation,
XX AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
XX disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
XX rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
XX thrombocytopenia, and ulcerative colitis. They may also be used to treat
XX complications of cancer, haemodialysis, extracorporeal circulation,
XX trauma and haematopoietic cancer, including lymphoma, leukaemia and
XX myeloma. Polynucleotides encoding HTFS proteins are useful for creating
XX transgenic animals to model human diseases, for diagnostic purposes and
XX to generate hybridisation probes useful in mapping the naturally
XX occurring genomic sequences. HTFS, and its catalytic or immunogenic
XX fragments are useful for screening libraries of compounds in a variety of
XX drug screening techniques. Antibodies which specifically bind HTFS may be
XX used for the diagnosis of disorders associated with the expression of
XX HTFS, or in assays to monitor patients being treated with HTFS or

CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.

Sequence 261 AA;

Query Match 9.0%; Score 25; DB 22; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFCLPTKKYADVILPRG 213
Db 187 TFWKPAFEFCLPTKKYADVILPRG 211

RESULT 12

AAB56582
ID AAB56582 standard; Protein; 337 AA.

AC AAB56582;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
XX vulnary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
XX antibacterial; gene therapy; neutral; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.

OS Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI (ROSE/) ROSEN C A.

PT Rosen CA, Ruben SM;

DR WPI; 2000-587513/55.

DR N-PSDB; AAF15785.

XX Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of

PT disorders such as prostate cancer -

XX Claim 11; Page 1566-1567; 2338pp; English.

XX AAF1566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytosolic,
XX cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX nephrotoxic, antiinfective, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention.

Sequence 337 AA;

Query Match 9.0%; Score 25; DB 21; Length 337;

Best Local Similarity 100.0%; Pred. No. 2.2e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFCLPTKKYADVILPRG 213
Db 263 TFWKPAFEFCLPTKKYADVILPRG 287

RESULT 13

ABP41393
ID ABP41393 standard; Protein; 337 AA.

AC ABP41393;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytosolic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 1p32.

OS Homo sapiens.

PN WO200206577-A1.

PD 03-JAN-2002.

PF 07-JUN-2001; 2001WO-US18569.

PR 07-JUN-2000; 2000US-209467P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

DR WPI; 2002-147878/19.

DR N-PSDB; ABQ54470.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g.

PT ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases -

XX Claim 11; SEQ ID NO 2525; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 337 AA;
 Query Match 9.0%; Score 25; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 2,2e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TPVKPAFEERFCLPTKXYADVIIPRG 213
 |||||
 Db 263 TPVKPAFEERFCLPTKXYADVIIPRG 287

RESULT 14
 AAE16594
 ID AAE16594 standard; Protein; 125 AA.
 XX
 AC AAE16594;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human 57658 protein prodomain consensus sequence #1.
 XX
 KM Human; uridine kinase-like protein; haematopoietic neoplastic disorder;
 KM 57658 protein; leukaemia; haematopoietic disorder; diabetes mellitus;
 KM rheumatoid arthritis; psoriasis; dermatitis; osteoporosis; rickets;
 KM sarcoma; myocardial infarction; hypertension; atherosclerosis;
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KM gene therapy; chromosome mapping; tissue typing; dermatological;
 KM cytostatic; osteopathic; cardiac; neuroprotective; nootropic;
 KM anticonvulsant; prodomain.
 KM
 OS Homo sapiens.
 XX
 PN WO200202761-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 28-JUN-2001; 2001WO-US21063.
 XX
 PR 30-JUN-2000; 2000US-216503P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA;
 XX
 PT WPI; 2002-140091/18.
 XX
 DR New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels -
 XX
 PS Disclosure; Fig 4; 103pp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays).

CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is the consensus sequence of the human 57658 protein
 CC prodomain, kinase uridine monophosphokinase transferase ATP-binding
 CC kinase-like ribonucleoside pyrimidine FIS cDNA domain.

SQ Sequence 125 AA;
 Query Match 5.4%; Score 15; DB 23; Length 125;
 Best Local Similarity 100.0%; Pred. No. 3.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 RSYKRTFSEPDHPG 260
 |||||
 Db 94 RSYKRTFSEPDHPG 108

RESULT 15
 ABB62307
 ID ABB62307 standard; Protein; 260 AA.
 XX
 AC ABB62307;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 13713.
 XX
 KM Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) BE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR N-PADB; ABL06410.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
 CC sequences (ABBS7737-ABBS7072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 260 AA;
 Query Match 5.4%; Score 15; DB 22; Length 260;
 Best Local Similarity 100.0%; Pred. No. 6.4e-06;

	Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	130		YPADVLFEGILVEY	144						
Db	135		YPADVLFEGILVEY	149						

Search completed: November 25, 2003, 08:09:03
Job time : 39 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:08:15 ; Search time 19 Seconds
(without alignments)
616.848 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDCESPAPEADRP...HPGLTSGKRSHESSSRPH 277

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB .pep:*
2: /cgn2_6/ptodata/2/1aa/5B COMB .pep:*
3: /cgn2_6/ptodata/2/1aa/6A COMB .pep:*
4: /cgn2_6/ptodata/2/1aa/6B COMB .pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB .pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1 .pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	93.9	260	4	US-09-536-647-2
2	45	16.2	277	4	US-09-536-647-3

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-2

Query Match 93.9%; Score 260; DB 4; Length 260;

Best Local Similarity 100.0%; Pred. No. 4.6e-252;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MASAGEDCESPAPEADRP...HPGLTSGKRSHESSSRPH 277
DB	1	MASAGEDCESPAPEADRP...HPGLTSGKRSHESSSRPH 277
QY	61	ILSODRFYKVLPAEQAKALKGOYNFDPDAFNDMHTLKNIVEGKIVEPTYPVTH 120
DB	61	ILSODRFYKVLPAEQAKALKGOYNFDPDAFNDMHTLKNIVEGKIVEPTYPVTH 120
QY	121	SRLEPTTVVYPADVVLFEGLIVFYSQEIIRDMFHLRLFVDTSDVRLSRRLDVRGRDL 180
DB	121	SRLEPTTVVYPADVVLFEGLIVFYSQEIIRDMFHLRLFVDTSDVRLSRRLDVRGRDL 180
QY	181	EQILTYTTPVPAPEEFCLPTKRYADVILIPRGVDMVAINLIVQHIDILNDICKMR 240
DB	181	EQILTYTTPVPAPEEFCLPTKRYADVILIPRGVDMVAINLIVQHIDILNDICKMR 240
QY	241	GGSGRSYKRTFSEPGDHPG 260
DB	241	GGSGRSYKRTFSEPGDHPG 260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Query Match 16.2%; Score 45; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 5.4e-37;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	100	TLKNIVEGKIVEPTYPVTHSRLEPTTVVYPADVVLFEGLIVFY 144
DB	100	TLKNIVEGKIVEPTYPVTHSRLEPTTVVYPADVVLFEGLIVFY 144

Search completed: November 25, 2003, 08:11:26
Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:09:10 ; Search time 31 seconds

(without alignments)
1648.090 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277

Sequence: 1 MASAGEDCESPAPADRP...HPGMLTSGKSHLESSSRPH 277

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 673684 seqs, 184443283 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: Published Applications AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	100.0	277	9	US-09-896-522-2
2	60	21.7	60	12	US-10-029-386-28765
3	25	9.0	337	10	US-09-925-300-1160
4	15	5.4	125	9	US-09-896-522-5

ALIGNMENTS

RESULT 1
US-09-896-522-2
Sequence 2, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Gluckemann, Maria A.

TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE OF INVENTION: USES THEREOF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 277
TYPE: PRT
ORGANISM: Homo sapiens
US-09-896-522-2

Query Match 100.0%; Score 277; DB 9; Length 277;
Best Local Similarity 100.0%; Pred. No. 4e-264;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAGEDCESPAPADRP...HPGMLTSGKSHLESSSRPH 60
DB 1 MASAGEDCESPAPADRP...HPGMLTSGKSHLESSSRPH 60
QY 61 ILSODRFYKVLTAEGKAKALNGQYNFDPDAFNDJMRITLKNIVEGKVEVPTDFVTH 120
DB 61 ILSODRFYKVLTAEGKAKALNGQYNFDPDAFNDJMRITLKNIVEGKVEVPTDFVTH 120
QY 121 SRLPETTVVYPADVVLFGGILVFYSQETIRDMHMLFVDVTSQVRLSRVLRDVRGRDL 180
DB 121 SRLPETTVVYPADVVLFGGILVFYSQETIRDMHMLFVDVTSQVRLSRVLRDVRGRDL 180
QY 181 EQLLYQYTFYVPAFEEPCLPKTYADVIIPRGVNMVAINIYOHIODILNGDICKMR 240
DB 181 EQLLYQYTFYVPAFEEPCLPKTYADVIIPRGVNMVAINIYOHIODILNGDICKMR 240
QY 241 GGSNGRSYKRTFSEPDHFGMLTSGKSHLESSSRPH 277
DB 241 GGSNGRSYKRTFSEPDHFGMLTSGKSHLESSSRPH 277

RESULT 2
US-10-029-386-28765
Sequence 28765, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GE
FILE REFERENCE: ABOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 28765
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR9.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Query Match 21.7%; Score 60; DB 12; Length 60;

Best Local Similarity 100.0%; Pred. No. 3.1e-51;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 VAINLIYQHIQDILNGDICKWHRGSGNSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 277
Db 1 VAINLIYQHIQDILNGDICKWHRGSGNSYKRTFSEPGDHPGMLTSGKSHLESSSRPH 60

RESULT 3

US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US2002015161A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-300-1160

Query Match 9.0%; Score 25; DB 10; Length 337;

Best Local Similarity 100.0%; Pred. No. 3.7e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 TFWKPAFEFCLPTKXYADVVIIPRG 213
Db 263 TFWKPAFEFCLPTKXYADVVIIPRG 287

RESULT 4

US-09-896-522-5
; Sequence 5, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Query Match 5.4%; Score 15; DB 9; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 246 RSYKRTFSEPGDHPG 260
Db 94 RSYKRTFSEPGDHPG 108

Search completed: November 25, 2003, 08:12:12
Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:07:30 ; Search time 19 Seconds
(without alignments)
1402.038 Million cell updates/sec

Title: US-09-896-522-2

Perfect score: 277

Sequence: 1 MASAGGDCESPAPFADRP.....HPGMLTSGKRSHESSSRPH 277

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : .PIR 76:*

1: pirl:*

2: pirl:*

3: pirl:*

4: pirl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description

No matches found

Search completed: November 25, 2003, 08:10:52
Job time : 19 secs

This Page Blank (uspto)


```

Db      121 SRLETTVVYPADVLFEGILVFSQELRDMFHLRLFPVDTSDVRLSRVLRVDRGRDL 180
Qy      181 EQLITQYTTTPYKPAPEEFCEPTKKYADVITPRGVDMVAIILYQHIQDILNGDICKWHR 240
Db      181 EQLITQYTTTPYKPAPEEFCEPTKKYADVITPRGVDMVAIILYQHIQDILNGDICKWHR 240
Qy      241 GGSNGRSYKRTFSEPGDHPGLTSGKRSHESSSRPH 277
Db      241 GGSNGRSYKRTFSEPGDHPGLTSGKRSHESSSRPH 277

RESULT 2
UCR1_MOUSE
ID      UCR1_MOUSE      STANDARD;      PRT;      277 AA.
AC      P52623;
DT      01-OCT-1996 (Rel. 34, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
DE      monophosphokinase 1) (Cytidine monophosphokinase 1).
GN      UCR1 OR UMPK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Salivary gland;
RX      MEDLINE=23388257; PubMed=12477932;
RA      Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datcchenko L., Maruina K., Farmer A.A., Rubin J., Hien L.,
RA      Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ueda T.B., Toshitsuki S., Carrinci P., Prange C.,
RA      Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA      Schermer A., Schein J.E., Jones S.J.M., Matra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [2]
RP      SEQUENCE OF 18-277 FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=97108719; PubMed=8951040;
RA      Ropp P.A., Traut T.W.;
RT      "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT      brain.";
RL      Arch. Biochem. Biophys. 336:105-112(1996).
CC      -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC      monophosphate and cytidine monophosphate. Does not phosphorylate
CC      deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC      as a phosphate donor. Can also phosphorylate cytidine and uridine
CC      nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC      thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC      benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC      methylcytidine, and N(4)-anisoylcytidine (By similarity).
CC      -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1- PATHWAY: Pyrimidine salvage pathway.
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
  
```

```

CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: BC025146; AAH25146.1; -.
DR      EMBL: L31783; AAB50568.1; -.
DR      MGI: 98904; Umpk.
DR      InterPro: IPR006083; PRK URK.
DR      InterPro: IPR00764; Uridine_kin.
DR      Pfam: PF00485; PRK_1.
DR      PRINTS: PR00988; URIDINKINASE.
DR      TIGRFAMs: TIGR00235; udk; 1.
FW      Transferase; Kinase; ATP-binding.
KW      NP_BIND 30
SQ      SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Query Match      16.2%; Score 45; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 2,9e-38;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      100 TLKNIVEKTEVEPTVDFTVTHSRLEPTTVVYPADVLFEGILVY 144
Db      100 TLKNIVEKTEVEPTVDFTVTHSRLEPTTVVYPADVLFEGILVY 144

RESULT 3
UCR2_HUMAN
ID      UCR2_HUMAN      STANDARD;      PRT;      261 AA.
AC      Q9B2X2; Q96RG5;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
DE      monophosphokinase 2) (Cytidine monophosphokinase 2).
GN      UCK2 OR UMPK.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX      MEDLINE=21203813; PubMed=11306702;
RA      Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;
RT      "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT      human uridine-cytidine kinases.";
RL      Mol. Pharmacol. 59:1181-1186(2001).
RN      [2]
RP      SEQUENCE OF 15-261 FROM N.A.
RC      TISSUE=Fibrosarcoma;
RX      MEDLINE=21385121; PubMed=11494055;
RA      Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RA      Fukushima M.;
RT      "Cloning and expression of uridine/cytidine kinase cDNA from human
RT      fibrosarcoma cells.";
RL      Int. J. Mol. Med. 8:273-278(2001).
CC      -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC      monophosphate and cytidine monophosphate. Does not phosphorylate
CC      deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC      as a phosphate donor. Can also phosphorylate cytidine and uridine
CC      nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC      thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC      benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC      methylcytidine, and N(4)-anisoylcytidine.
CC      -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1- PATHWAY: Pyrimidine salvage pathway.
CC      -1- TISSUE SPECIFICITY: Expressed in placenta.
CC      -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
  
```


CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF236637; AKL14053.1; -
 CC EMBL; AB062451; BAB56162.1; -
 CC Gene; HGNC:12562; UMPK.
 CC InterPro; IPR006082; PRK.
 CC InterPro; IPR006083; PRK_URK.
 CC InterPro; IPR000764; Uridine_kin.
 CC Pfam; PF00485; PRK.1.
 CC PRINTS; PR00478; PHRIBLKINASE.
 CC PRINTS; PR00988; URIDINKINASE.
 CC TIGRFAme; TIGR00235; udk.1.
 CC Transferase; Kinase; ATP-binding.
 CC NP_BIND 27 34 ATP (POTENTIAL).
 CC SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;
 SQ

Query Match 9.0%; Score 25; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 8.5e-18;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECUPTKKYADVITPRG 213
 |||||
 DB 187 TFWKPAFEFECUPTKKYADVITPRG 211

RESULT 4
 UCK2 MOUSE STANDARD; PRT; 261 AA.
 AC 099PM9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
 monophosphokinase 2) (Cytidine monophosphokinase 2).
 GN UCK2 OR UMPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21203813; PubMed=11306702;
 RA Van Rompay A.R., Norda A., Lindén K., Johansson M., Karlsson A.;
 RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
 human uridine-cytidine kinases.";
 RL Mol. Pharmacol. 59:1181-1186(2001).
 CC -|- FUNCTION: Phosphorylates uridine and cytidine to uridine
 monophosphate and cytidine monophosphate. Does not phosphorylate
 deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
 as a phosphate donor. Can also phosphorylate cytidine and uridine
 nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
 thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
 benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
 methylcytidine, and N(4)-amsoylcytidine (By similarity).
 CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
 CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
 CC -|- PATHWAY: Pyrimidine salvage pathway.
 CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF236636; AKL14052.1; -
 CC GMD; MG1:1931744; UCK2.
 CC InterPro; IPR006082; PRK.

DR InterPro; IPR006083; PRK_URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAme; TIGR00235; udk.1.
 CC Transferase; Kinase; ATP-binding.
 CC NP_BIND 27 34 ATP (POTENTIAL).
 CC SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;
 SQ

Query Match 9.0%; Score 25; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 8.5e-18;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFWKPAFEFECUPTKKYADVITPRG 213
 |||||
 DB 187 TFWKPAFEFECUPTKKYADVITPRG 211

RESULT 5
 UCK DROME STANDARD; PRT; 260 AA.
 ID UCK_DROME
 AC 09VC99;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
 monophosphokinase) (Cytidine monophosphokinase).
 GN CG6364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhargava D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downe M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,
 RA Hostein D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Idaliti M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mozhayeva A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein-Volhard P., Patel J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskae R., Tector A.C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

```

RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000).
CC  -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC  -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC  -1- PATHWAY: Pyrimidine salvage pathway.
CC  -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE003747; AAF56274.1; -.
DR  FlyBase; FBgn0039179; CG6364.
DR  InterPro; IPR006083; PRK_UK.
DR  InterPro; IPR000764; Uridine_kin.
DR  Pfam; PF00485; PRK_1.
DR  PRINTS; PR00988; URIDINKINASE.
KW  Hypothetical protein; Transferase; Kinase; ATP-binding.
FT  NP_BIND 34 41 ATP (POTENTIAL).
SQ  SEQUENCE 260 AA; 29213 MM; 285742370492566 CRC64;

```

```

Query Match      5.4%; Score 15; DB 1; Length 260;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  130 YPADVLFEGILVY 144
    |||||
Db  135 YPADVLFEGILVY 149

```

Search completed: November 25, 2003, 08:09:29
 Job time : 11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2003, 08:07:00 ; Search time 35 Seconds
(without alignments)
2042.302 Million cell updates/sec

Title: US-09-896-522-2
Perfect score: 277
Sequence: 1 MASAGEDCESPAPADRPH.....HPGMLTSGKRSKSHLESSSRPH 277

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 segs, 258052604 residues

Word size : 15

Total number of hits satisfying chosen parameters: 6

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_virus:*
15: sp_bacteriap:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	61.0	201	4 Q96B0	Q96B0 homo sapien
2	25	9.0	105	11 Q9QY8	Q9QY8 rattus norv
3	25	9.0	111	4 Q9BU42	Q9BU42 homo sapien
4	25	9.0	261	11 Q8C476	Q8C476 mus musculi
5	15	5.4	111	4 Q92528	Q92528 homo sapien
6	15	5.4	260	5 Q8MRJ1	Q8MRJ1 drosophila

ALIGNMENTS

RESULT 1
Q96B0 PRELIMINARY; PRT; 201 AA.
AC Q96B0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015547; AA015547.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
SQ SEQUENCE 201 AA; 22790 MW; 0E5F2F00F7B5363 CRC64;

Query Match 61.0%; Score 169; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.1e-170; Indels 0; Gaps 0;
Matches 169; Conservative 0; Mismatches 0;

QY 1 MASAGEDCESPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEOROKVY 60
DB 1 MASAGEDCESPAPADRPORPFLIGVSGTASGKSTVCEKIMELLGONEVEOROKVY 60
QY 61 IISQDRFYKVLTAEOAKALKQYNFDPDAPNDLMHRTLNIVGKTVETVYDFVTH 120
DB 61 IISQDRFYKVLTAEOAKALKQYNFDPDAPNDLMHRTLNIVGKTVETVYDFVTH 120
QY 121 SRLETTVVYPADVVFEGILVFYSGEIRDMHRLRFVYDSDVRLSR 169
DB 121 SRLETTVVYPADVVFEGILVFYSGEIRDMHRLRFVYDSDVRLSR 169

RESULT 2
Q9QY8 PRELIMINARY; PRT; 105 AA.
AC Q9QY8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, last annotation update)
DE Uridine kinase (Fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yao T., Watanabe S., Okajima S., Hirasawa Y., Fushiki S.;
RT "Up-regulated uridine kinase gene identified by RLCs in the ventral
horn after crush injury to rat sciatic nerves."
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BA083085.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11967 MW; 9DB43C15E07EB029 CRC64;

Query Match 9.0%; Score 25; DB 11; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.6e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 TFVKPAFEFFCLPTKKYADVILIRG 213

Db 36 TFWKPAFEFCLPTKXADVIIPRG 60

RESULT 3

Q9BU42 PRELIMINARY; PRT; 111 AA.
AC Q9BU42; 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Similar to uridine monophosphate kinase.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strauberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002906; AA02906.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
KW Kinase.
SQ SEQUENCE 111 AA; 12587 MW; E6688B1B86FA32A9 CRC64;

Query Match 9.0%; Score 25; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 4.8e-18;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 TFWKPAFEFCLPTKXADVIIPRG 213
Db 37 TFWKPAFEFCLPTKXADVIIPRG 61

RESULT 4

Q8C476 PRELIMINARY; PRT; 261 AA.
AC Q8C476; 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Uridine-cytidine kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573 (2002).
RL Nature 420:563-573 (2002).
DR EMBL; AK082837; BAC38646.1; -
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Qy 189 TFWKPAFEFCLPTKXADVIIPRG 213
Db 187 TFWKPAFEFCLPTKXADVIIPRG 211

RESULT 5

Q92528

ID Q92528 PRELIMINARY; PRT; 111 AA.

AC Q92528; 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 5'-terminal region of UMK.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96411689; PubMed=8912458;
RA Oaki K., Kuraki T., Hayashi S., Nakamura Y.;
RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
RT a differential mRNA display method.";
RL Genomics 36:316-319 (1996).
DR EMBL; D78335; BA011349.1; -
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Query Match 5.4%; Score 15; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 TFWKPAFEFCLPTK 203
Db 37 TFWKPAFEFCLPTK 51

RESULT 6

Q8MRJ1 PRELIMINARY; PRT; 260 AA.
AC Q8MRJ1; 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ID13909P.
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Pargass V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY119583; AAM50237.1; -
DR FlyBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK.1.
DR PRINTS; PR00988; URIDINKINASE.
SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16E5 CRC64;

Query Match 5.4%; Score 15; DB 5; Length 260;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 130 YPADVLFEGILVFY 144
Db 135 YPADVLFEGILVFY 149

Wed Nov 26 18:16:18 2003

us-09-896-522-2.01.rsp

Page 3

Search completed: November 25, 2003, 08:10:18
Job time : 35 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:50:54 ; Search time 17.5085 Seconds
(without alignments)
7849.063 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 536
Sequence: 1 gfggggctgcctccgacaccc.....ccaggtctgfggggacag 1624

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Word size: 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/rnat_21112003_184146_3259/app_query.fasta.1.2830
-DB=Issued Patents AA -QFMT=fasta -SUFFIX=n2p_ol1.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.QCEN_1.1.29@rnat_21112003_184146_3259 -NCPV=6 -ICPV=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	48.5	260	4	US-09-536-647-2
2	45	8.4	277	4	US-09-536-647-3

ALIGNMENTS

RESULT 1
US-09-536-647-2
; Sequence 2, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen

APPLICANT: Johnson, Randall
TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
FILE REFERENCE: GP50020
CURRENT APPLICATION NUMBER: US/09/536,647
CURRENT FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 260
TYPE: PRT
ORGANISM: Human
US-09-536-647-2

Alignment Scores:
Pred. No.: 1.74e-240
Score: 260.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 48.51%
DB: 4
Length: 260
Matches: 260
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-896-522-1 (1-1624) x US-09-536-647-2 (1-260)

QY	94	ATGCGTTCCGCGGAGGCGAAGACTGCGAGACCCCGCGGAGCGGACCGCTCGCAC	153
DB	1	MetAlaserAlaglygylglnaspcysglsuserProAlaProGlnAlasparProHis	20
QY	154	CAGCGCCCTTCTGTATAGGGGTAGCGCGGCACTCCAGCGGGAAGTGCACCTGTGT	213
DB	21	GlnArgProPheLeuilegylValsergylThrAlaserGlyLysSerThrValCys	40
QY	214	GAGAAGATCATGAGTGTCTGGGACAGAACGAGTGGAAACAGCGGACCGGAAGTGTCT	273
DB	41	GlnLysIleMetGlnLeuLeuGlnGlnasnglnValGlnGlnArgGlnArgLysValVal	60
QY	274	ATCTCGAGCCAGAGACAGTTCTTCAAGGCTCTGACGCGGACAGACAGGCGCTTG	333
DB	61	IleLeuSerGlnAspArgPheItyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu	80
QY	334	AAAGACAGTACATTTTGGACCATCCAGATGCGCTTGTATATGATTTGTGACAGACT	393
DB	81	LysGlyGlnItyrAspPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr	100
QY	394	CTGAAGAATATGCTGAGGGCAAAACGTTGAGGTGCCGACTATGATTTTGTGACAC	453
DB	101	LeuLysAsnIleValGlnGlnItyrThrValGlnValProThrItyrAspPheValThrHis	120
QY	454	TCAGGTTACAGAGACACGCGTGTCTTACCTGCGGACGCTGTCTGTGAGGCGATC	513
DB	121	SerArgLeuProGlnThrThrValItyrProAlaAspValValLeuPheGlnGlyIle	140
QY	514	TTGGTGTCTTACAGCGGAGATCCGGAGATGTTCCACTTGGCTCTTGTGACACC	573
DB	141	LeuValPheItyrSerGlnGlnIleItyrAspMetPheHisLeuArgLeuPheValAspThr	160
QY	574	GACTCCGACGTACGCTGTCTTGAAGATTTCCGGAGCGTCCGCCGAGGAGGACCTG	633
DB	161	AspSerAspValArgLeuSerThrArgValLeuArgAspValArgArgGlyItyrAspLeu	180
QY	634	GAGCAGATCTGACGACAGTACACACCTTGGAGACCGGCTTGGAGAGTTTGGCGTG	693
DB	181	GlnGlnIleLeuThrGlnItyrThrThrPheValLysProAlaPheGlnGlnItyrCysLeu	200
QY	694	CCGACAAAGAGATGATCCGATGTGATCATCCAGAGAGTGGACAAATATGTTGCCATC	753
DB	201	ProThrLysItyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
QY	754	AACCTGATGTCACACATCCAGACATTTGGAATGTGGAATTTGCAAAATGGACCGCA	813
DB	221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysItyrHisArg	240
QY	814	GGAGGTCCAATGGCGGAGCTTACAGCGGACCTTTTGTGACCGGAGGACCACTGGAG	873

Db 241 GlycylserArgGlyArgSerTyrIyArgThrPheSerGluProGlyAspHisProGly 260

RESULT 2

US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708e1 Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; CURRENT FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:

Pred. No.:	1.38e-34	Length:	277
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.40%	Indels:	0
DB:	4	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-536-647-3 (1-277)

QY	391	ACTCTGAGAACATCGTGGAGGGCAAAACGGTGGAGTGCCGACCTATGATTTTGACA	450
Db	100	ThrLeuYsaenIleValGluGlyLysThrValGluValProThrTyrAspPheValThr	119
QY	451	CACTCAAGGTACCAAGACGACGAGTGTACCCCTGGCGACGTGTGTTGAGGGC	510
Db	120	HisSerArgLeuProGluThrThrValValTyrProIleAspValValLeuPheGluGly	139
QY	511	ATCTGGTGTCTAC	525
Db	140	IleLeuValPheTyr	144

Search completed: November 25, 2003, 08:01:00
Job time : 19.5085 secs

Db 161 AsperasperaValaGlueSerArgValaLeuArgAspValaGlyArgAspLeu 180
Qy 634 GAGCAGATTGTGAGCAGTACACACCTTCGTGAAGCCGCTTCGAGAGTTCTGCTG 693
Db 181 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 200
Qy 694 CCGACAAAGAGTATGCCGATGTATCATCCACAGAGAGTGGACATATGTTGCCATC 753
Db 201 ProThrLysLysTyrAlaAspValaIleIleProArgGlyValaAspAspMetValaIle 220
Qy 754 AACCTGATCGTGCAGCATCATCAGACATTTGATGATGTGACATCTGCAATGAGCAGCA 813
Db 221 AsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTyrPheLysArg 240
Qy 814 GAGAGGTCCATGGGCGGAGCTACAGCGACCTTTTCTGAGCGAGGAGACACCTCTGG 873
Db 241 GlyLysArgGlnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 260
Qy 874 ATGCTGACCTTGGCAACGCTCATTTGAGTCCAGCGACGACCCAC 924
Db 261 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 277

RESULT 2

US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALU8 8.00e-26
US-10-029-386-28765

Alignment Scores:

Pred. No.:	1.85e-47	Length:	60
Score:	60.00	Matches:	60
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.19%	Indels:	0
DB:	12	Gaps:	0

US-09-896-522-1 (1-1624) x US-10-029-386-28765 (1-60)

Qy 745 GTTGCCATCAACTGATCGTACAGACATCCAGACATTTGATGGTGAATCGCAAA 804
Db 1 ValaIleAsnLeuIleValaGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 20
Qy 805 TGGCAGCGAGAGGGTCCATGGGCGGAGCTACAGCGGACCTTTTCTGAGCGAGGAG 864
Db 21 TrpHisArgGlyLysArgSerHisLysArgSerTyrLysArgThrPheSerGluProGlyAsp 40

Qy 865 CACCTGGAGTGTGACTCTTGCGCAACGCTCATTTGAGTGCAGACAGACCCAC 924
Db 41 HisProGlyMetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 60

RESULT 3

US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; APPLICANT: PA101
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160

Alignment Scores:

Pred. No.:	1.76e-14	Length:	337
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.66%	Indels:	0
DB:	10	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-925-300-1160 (1-337)

Qy 658 ACCTTGTTGAAAGCGGCTTCGAGAGTCTGCTGCCGCAAAAGATATGCCGATGTG 717
Db 263 ThrPheValLysProAlaPheGluGluPheCysLeuProThrLysTyrAlaAspVal 282

Qy 718 ATCATCCACAGAGA 732
Db 283 IleIleProArgGly 287

RESULT 4

US-09-764-877-1422
; Sequence 1422, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1422
; LENGTH: 73

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (30)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (72)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1422

Alignment Scores:

Pred. No.:	9.87e-09	Length:	73
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.54%	Indels:	0
DB:	10	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-764-877-1422 (1-73)

QY 1521 CCACTGAGGAGATTTTCCCCCTTGCTTGCCCTTAAGGCACAGAGGAGCGG 1577

Db 11 ProleuArgAspValPheProCysLeuCyLeuLysGlyArgGluAaArg 29

RESULT 5

US-09-896-522-5

Sequence 5, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Gluckemann, Maria A.
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
TITLE OF INVENTION: USRS THEREOF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US/09/896,522
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 125
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-5

Alignment Scores:

Pred. No.:	5.57e-05	Length:	125
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.80%	Indels:	0
DB:	9	Gaps:	0

US-09-896-522-1 (1-1624) x US-09-896-522-5 (1-125)

QY 829 CCGAGCTACAGCGACCTTTTCTGAGCCAGGCGACACCCCTGG 873
Db 94 ArgSerTyrLysArgThrPheSerGluProGlyAspHisProGly 108

Search completed: November 25, 2003, 08:04:07
Job time : 58.8202 secs

This Page Blank (uspto)


```

FEATURES
  source
    location/Qualifiers
      1. .1118
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5532885"
        /issue_type="leiomysarcoma"
        /lab_host="DH10B (phage-resistance)"
        /clone_id="NH_MGC_71"
        /note="Organ: uterus; Vector: pCMVSPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 2.1 kb."
BASE COUNT      234 a      311 c      361 g      209 t      3 others
ORIGIN
Query Match      54.9%; Score 890.8; DB 12; Length 1118;
Best Local Similarity 90.9%; Pred. No. 3.6e-179;
Matches 1005; Conservative 0; Mismatches 89; Indels 12; Gaps 5;

QY 49 GAAGGGCGGGCGGGGACCCGATGCGCGGAGCCGAGCCGACCGGAGCTTCGGCGGGA 108
DB 2 GGAGCGCTGGCGGGGACCCGATGCGCGGAGCCGAGCCGAGCTTCGGCGGGA 61

QY 109 GGCAGAACTGCGAGAGCCCGCGCGGAGCCGACCGCTCCGACAGCGGCGCTTCCTG 168
DB 62 GCGCAAGATCGCGAGAGCCCGCGCGGAGCCGACCGCTCCGACAGCGGCGCTTCCTG 121

QY 169 ATAGGGGTAGCGCGGCACTGCGAGCGGAGTGA CCGTGTGTGAGAAAGATCATGAG 228
DB 122 ATAGGGGTAGCGCGGCACTGCGAGCGGAGTGA CCGTGTGTGAGAAAGATCATGAG 181

QY 229 TTGCTGGGACAGACAGGTGGAACAGCGGAGCGGAGGTGATCTCGACAGAGAC 288
DB 182 TTGCTGGGACAGACAGGTGGAACAGCGGAGCGGAGGTGATCTCGACAGAGAC 241

QY 289 AGGTTCTACAAGGTCTGACGCGGACAGAGAGCCAGAGCGCTTGAAGAAGACAGTAAAT 348
DB 242 AGGTTCTACAAGGTCTGACGCGGACAGAGAGCCAGAGCGCTTGAAGAAGACAGTAAAT 301

QY 349 TTGACCATCCAGATGCTTGTATGATTTGATGCA CAGAGACTTGAAAGACATCGTG 408
DB 302 TTGACCATCCAGATGCTTGTATGATTTGATGCA CAGAGACTTGAAAGACATCGTG 361

QY 409 GAGGGCAAAACGGGAGGCGGACCTATGATTTTGTGACACCTGAAGTTACCGAG 468
DB 362 GAGGGCAAAACGGGAGGCGGACCTATGATTTTGTGACACCTGAAGTTACCGAG 421

QY 469 ACCAGGTGCTCTACCTGCGGAGGTGTTCTGTTGAGGAGCATTTGTTCTACAGC 528
DB 422 ACCAGGTGCTCTACCTGCGGAGGTGTTCTGTTGAGGAGCATTTGTTCTACAGC 481

QY 529 CAGAGATCCGGGACATGTTCCACTCGGCGCTCTTCTGTGACACCGACTCCGAGTCAAG 588
DB 482 CAGAGATCCGGGACATGTTCCACTCGGCGCTCTTCTGTGACACCGACTCCGAGTCAAG 541

QY 589 CTGCTGGAAGAGTCTCCGGGAGTGCAGCGGAGGAGGACCTGCGGAGCATTTCTACAG 648
DB 542 CTGCTGGAAGAGTCTCCGGGAGTGCAGCGGAGGAGGACCTGCGGAGCATTTCTACAG 601

QY 649 CAGTACACCACTTCTGTAAGCCGCGCTTGCAGAGAGTTCTGCTGCGGACAAAGAATAT 708
DB 602 CAGTACACCACTTCTGTAAGCCGCGCTTGCAGAGAGTTCTGCTGCGGACAAAGAATAT 661

QY 709 GCCGATGTATCATCCACGAGAGTGAACAATATGTTGTCATCAACTGATGTGAG 768
DB 662 GCCGATGTATCATCCACGAGAGTGAACAATATGTTGTCATCAACTGATGTGAG 721

QY 769 CACATCAGAGCATTTGAAATGTGACATCTGCAATGAGACCGAAGAGGTCCAAATGG 828
DB 722 CACATCAGAGCATTTGAAATGTGACATCTGCAATGAGACCGAAGAGGTCCAAATGG 781

QY 829 CGGAGCTACAAGCGGAGCTTTCTGAGCAAGGAGGAGCAACCTGAGTGTGACCTTGG 887

```

```

DB 782 CGAGGTACAAGCGACCTTTTCTGAGCGAGGAGACCACTGNGATGTGACTCTG 841
QY 888 CAACGGTCACTTTGAGTCCAGACGACA -CCCCACTGAGGGGCTGCCAGCTCAG 946
DB 842 CAACGGTCACTTTGAGTCCAGACGACA -CCCCACTGAGGGGCTGCCAGCTCAG 901
QY 947 GCAGGTCTCCCGCGG---CATGTGTTCAAGGACCTGAGGCGGCGGACCCAC 1003
DB 902 GCAGGTCTCCCGCGGCGGCGGATGGGGGTTTCGAACTGAACCTGGGAAAGCCACCA 961
QY 1004 ACCCACTGCTCTCTCTGCGGCGACCCC---AGGGAGTGTGACAGAGGCTTCCTGA 1060
DB 962 ACCCACTGCTCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1021
QY 1061 CTCAGAGTGAAGAACTCAGATGTGATC---AATCTCACTTCTGAGACCTGA 1116
DB 1022 CTCGAGGAGGCGGAACTCAGATGTGATC---AATCTCACTTCTGAGACCTGA 1081
QY 1117 GCGCTTCTGAGGTTTCAAGCACTT 1142
DB 1082 GCGCTTCTGAGGTTTCAAGCACTT 1107

RESULT 2
BX343101
LOCUS BX343101 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
ACCESSION BX343101
VERSION BX343101.1 GI:30334170
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1142)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr
library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DL009YD15&cluster=4968.r. Contact :
Peng Liang Email: filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope Sequence ID : CS0DL009YD15
FEATURES
  source
    location/Qualifiers
      1. .1114
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DL009YD15"
        /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
        /cell_line="RAMOS CELL LINE"
        /clone_id="Homo sapiens B CELLS (RAMOS CELL LINE) COT
        25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
        primer. Five prime end enriched, double-strand cDNA was
        digested with Not I and cloned into the Not I and EcoR V
        sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      247 a      291 c      360 g      202 t      14 others
ORIGIN
Query Match      52.9%; Score 859; DB 13; Length 1114;
Best Local Similarity 92.8%; Pred. No. 2.1e-172;
Matches 966; Conservative 8; Mismatches 60; Indels 7; Gaps 7;
6 GTGCTCCGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 65

```

```

72 GTGCGCTCCGACCTCGGCGCTGCGCGCGCGCGCGCTGGGGAAGGGCGCGCGG
QY 66 GACCCGATGCGCGGAGCGGAGCCGAGATGCTTCGCGGAGGCGAACACTGCGAG
Db 132 GACCCGATGCGCGGAG-8GAGGCGGAGATGCTTCGCGGAGGCGAACACTGCGAG
QY 126 CCGCGCGCGGAGCGGAGCGGCGCGCGCGCTTCGATAGAGGGTGAAGCGCGG
Db 191 CCGCGCGCGGAG-8GAGGCGGAGCGGCGCGCGCGCTTCGATAGAGGGTGAAGCGCGG
QY 186 CACTGCGAGCGGAGATGCGACCGTGTGTGAGAGATGATGAGTGTGCGAGAGCA
Db 250 CACTGCGAGCGGAGATGCGACCGTGTGTGAGAGATGATGAGTGTGCGAGAGCA
QY 246 GGTGGAACAGCGGAGCGGAGAGTGTGATCTGAGCGAGAGAGTTCACAAAGTCT
Db 310 GGTGGAACAGCGGAGCGGAGAGTGTGATCTGAGCGAGAGAGTTCACAAAGTCT
QY 306 GACGCGAGAGAGAGCGGAGCGGAGAGAGAGTTCACAAAGTTCACAAAGTTC
Db 370 GACGCGAGAGAGAGCGGAGCGGAGAGAGAGTTCACAAAGTTCACAAAGTTC
QY 366 CTTGATTAATGATTTGATGACAGACTCTGAGAGAGATGATGAGAGAGTTC
Db 430 CTTGATTAATGATTTGATGACAGACTCTGAGAGAGATGATGAGAGAGTTC
QY 426 GGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 490 GGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 486 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 550 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 546 GTTCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 610 GTTCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 606 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 670 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 666 GAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 730 GAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 726 ACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 790 ACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 786 GAATGATGATCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 850 GAATGATGATCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 846 CTTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 909 CTTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 906 GTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 968 GT-CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 966 TGTGTGTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 1025 TGTGTGTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1026 ACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db 1085 KTAGMRCGAGCTTCTTATCAG 1105

```

RESULT 3

```

BX372390/c 941 bp mRNA linear EST 08-MAY-2003
LOCUS BX372390 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL009YD15 3-PRIME, mRNA sequence.
ACCESSION BX372390
VERSION BX372390.1 GI:30456111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 941)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1046ZH05_CS04360_1cluster=4968.r.
Contact: Feng Liang Email: fliang@life.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID: CS0BA1046ZH05_CS04360_1.
FEATURES
    source
        1..941
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DL009YD15"
            /cell_line="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
            /cell_line="RAMOS CELL LINE"
            /clone_1ib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
            25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and cloned into the Not I and EcoR V
            sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 196 a 277 c 251 g 217 t
ORIGIN
Query Match 51.8%; Score 841.2; DB 13; Length 941;
Best Local Similarity 96.8%; Pred. No. 1.3e-168;
Matches 912; Conservative 0; Mismatches 23; Indels 7; Gaps 5;
QY 339 ACAGTACATTTTGAACCATCCAGATGCTTGTATATATTTGATGACAGAGCTGTGAA 398
Db 940 ACAGTACATTTTGAACCAACAAATGCTTGTATATATTTGATGACAGAGCTTTGAA 881
QY 399 GAACATGTCGAGAGGCAAAACGGTGA-8GTGCGGACATGATTTTGTGACACACTCAA 457
Db 880 AACCATGTCGAGAGGCAAAACGGTGAAGGTCGCTTATATATTTTGAACCCATCAA 821
QY 458 GGT--ACCAAGACCAAGTGTCTTACCTTC-8GAGCTGTTCTGTTAGAGGATCT 514
Db 820 AGGTACACAGAGAGACCAAGTGTCTTACCTTCGAGGACCTGTTGAGAGGATCT 761
QY 515 TGTGTTTACAGCAGAGAGATCCGGG-ACATGTTCACTGAGGCTTCTGAGACAC 573
Db 760 TGTGTTTACAGCAGAGAGATCCGGGAACATGTTCACTGAGGCTTCTGAGACAC 701
QY 574 GACTCCGAGTGAAGCTGTCTGAAAGAGTTCGCGGAGAGTGCAGGAGAGAGAGCTG 633
Db 700 GACTCCGAGTGAAGCTGTCTGAAAGAGTTCGCGGAGAGTGCAGGAGAGAGAGCTG 641
QY 634 GAGCAGATTTGACGAGATGACACACTTCTGTAAGCCGCTTGAAGAGATTTGCTGCTG 693
Db 640 GAGCAGATTTGACGAGATGACACACTTCTGTAAGCCGCTTGAAGAGATTTGCTGCTG 581
QY 694 CCGACAAAGATTAAGCCAGTGTATCATCCAGAGAGAGTGAACATATGTTGCCATC 753

```

```

Db      580 CCGGCAAAAGATGTCCGATGTGATCATCCACGAGGAGTGACATATGTTGCCATC 521
Qy      754 AACCTGATCTGTGACGACATCATCAGACATTTCTGAATGGTGACATCTGCATAATGGCACCGA 813
Db      520 AACCTGATCTGTGACGACATCATCAGACATTTCTGAATGGTGACATCTGCATAATGGCACCGA 461
Qy      814 GAGAGGTCCTAATGGGCGGAGCTAAGCGGACCTTTTCTGAGCGAGGGGACCACTCGGG 873
Db      460 GAGAGGTCCTAATGGGCGGAGCTAAGCGGACCTTTTCTGAGCGAGGGGACCACTCGGG 401
Qy      874 ATGTGACCTCTGGCAAAACGGTCACTTTGGAGTCCAGACAGACACCCACTGAGGGCT 933
Db      400 ATGTGACCTCTGGCAAAACGGTCACTTTGGAGTCCAGACAGACACCCACTGAGGGCT 341
Qy      934 GCCGAGCTCTAGGCGGAGGCTCTCCCGCCGCGATGTGTTCAGGAGCTGAGCTGGGAC 993
Db      340 GCCGAGCTCTAGGCGGAGGCTCTCCCGCCGCGATGTGTTCAGGAGCTGAGCTGGGAC 281
Qy      994 GCCGAGCTCTAGGCGGAGGCTCTCTCTGCGGACACCCAGGGAGTGTAGAGGAGGCC 1053
Db      280 GCCGAGCTCTAGGCGGAGGCTCTCTCTGCGGACACCCAGGGAGTGTAGAGGAGGCC 221
Qy      1054 TTCTCTACTGAGAGTGAAGAACTGAGATGTGTCTACTGAGACTCACTTGTGAGCACTG 1113
Db      220 TTCTCTACTGAGAGTGAAGAACTGAGATGTGTCTACTGAGACTCACTTGTGAGCACTG 161
Qy      1114 ACAGCGCTTCTGAGGTTTTCAGGCTTCTGAGGCTGCTGGTTTAAAGATCCTCTAGG 1173
Db      160 ACAGCGCTTCTGAGGTTTTCAGGCTTCTGAGGCTGCTGGTTTAAAGATCCTCTAGG 101
Qy      1174 TCATCTGAGAAATGACAGAAATGTGACAGAAAGCTGGGAGGCTTCTGAGGAAATGTAG 1233
Db      100 TCATCTGAGAAATGACAGAAATGTGACAGAAAGCTGGGAGGCTTCTGAGGAAATGTAG 41
Qy      1234 GCACATTTATGGGAAATTTGAGAGAGAGCCTTAGACACTGGC 1275
Db      40 GCACATTTAT--GGGAAATTTAGAGAGAGCCTTAGACACTGGC 1

RESULT 4
LOCUS   BG826894
DEFINITION 60275097BP1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903819 5',
mRNA sequence.
ACCESSION BG826894
VERSION   BG826894.1
KEYWORDS  GI:14174481
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 916)
          NIH-MGC http://mgs.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: rgs@bgl-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Ling Hong/Rubin Laboratory
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LNCM802 row: k column: 20
          High quality sequence stop: 843.
          Location/Qualifiers
            1..916
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4903819"

```

```

/issue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_17"
/notes="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 203 a 239 c 309 g 165 t

ORIGIN

Query Match 50.6%; Score 821.2; DB 12; Length 916;
Best Local Similarity 97.8%; Pred. No. 2,3e-164;
Matches 896; Conservative 0; Mismatches 13; Indels 7; Gaps 6;

```

Qy      15 GACCTCGGCGCTGGGCGGGGCGCGCGGCGGGAAGGGGCGGGCGGGAGCCCGATG 74
Db      1 GACCTCGGCGCTGGGCGGGGCGCGCGGCGCTGGGGAAGGGGCGGGAGCCCGATG 60
Qy      75 CGGAGGAGCGGAGGCGGAGATGGCTTTGCGCGGAGGCGAAGACTGCGAGACCCCGCGC 134
Db      61 CGGAGGAGCGGAGGCGGAGATGGCTTTGCGGCGGAGGCGAAGACTGCGAGACCCCGCGC 120
Qy      135 GAGGCGGACCGTCCGCAACAGGCGGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAG 194
Db      121 GAGGCGGACCGTCCGCAACAGGCGGCTTCTGATAGGGGTGAGCGGCGGCACTGCGAG 180
Qy      195 CGGAGGAGCGACCGTGTGTGAGAGATCAATGAGTGTCTGGGACAGAGAGGTGAGAA 254
Db      181 CGGAGGAGCGACCGTGTGTGAGAGATCAATGAGTGTCTGGGACAGAGAGGTGAGAA 240
Qy      255 GCGGAGCGGAGAGGTGTCTATCTGAGCGGAGGAGGAGGCTTCTAGCGGCA 314
Db      241 GCGGAGCGGAGAGGTGTCTATCTGAGCGGAGGAGGAGGCTTCTAGCGGCA 300
Qy      315 GCAGAGGCGGAGGCGGCTTGAAGAGAGTAAATTTTGAACATCCAGATGCCCTTATTA 374
Db      301 GCAGAGGCGGAGGCGGCTTGAAGAGAGTAAATTTTGAACATCCAGATGCCCTTATTA 360
Qy      375 TGATTGATGACAGAGACTCTGAAAGACATCTGAGAGGCGAAACCGTGGAGGTGCCGAG 434
Db      361 TGATTGATGACAGAGACTCTGAAAGACATCTGAGAGGCGAAACCGTGGAGGTGCCGAG 420
Qy      435 CTATGATTTTGTGACACACTCAAGGTTTACAGAGACCGGTGTCTACCTCGGAGCGT 494
Db      421 CTATGATTTTGTGACACACTCAAGGTTTACAGAGACCGGTGTCTACCTCGGAGCGT 480
Qy      495 GGTTCGTGTTGAGGAGATCTGTGTCTTACAGCGGAGAGATCCGGGAGATGTTCCAGCT 554
Db      481 GGTTCGTGTTGAGGAGATCTGTGTCTTACAGCGGAGAGATCCGGGAGATGTTCCAGCT 540
Qy      555 GCGCCTCTTCTGAGAGACGAGACTCCGAGCTCAGAGGCTGTCTCGAAGAGTTCGCGGAGCT 614
Db      541 GCGCCTCTTCTGAGAGACGAGACTCCGAGCTCAGAGGCTGTCTCGAAGAGTTCGCGGAGCT 600
Qy      615 GCGCGAGGAGGAGGAGCTGTGAGAGATTTGAGAGCACTACACCACTTGTGAGAGCGGC 674
Db      601 GCGCGAGGAGGAGGAGCTGTGAGAGATTTGAGAGCACTACACCACTTGTGAGAGCGGC 660
Qy      675 CTTCGAGGAGTTCGTGCTGCGGACAAAGAAAGTATGCGGATGATATCCAGAGAGAGT 734
Db      661 CTTCGAGGAGTTCGTGCTGCGGACAAAGAAAGTATGCGGATGATATCCAGAGAGAGT 720
Qy      735 GAGCAATA--TGGTTCATCAACCTGATGTGAGAGACA--TCCAGAGACATTC--TGATGG 791
Db      721 GAGCAATATTTGGTTCATCAACCTGATGTGAGAGACA--TCCAGAGACATTC--TGATGG 780
Qy      792 TGAATCTGCAATATGAGACCGAGAGAGGATCAAT--GAGCGGAGCTTCAAG--CGACCTT 848
Db      781 TGAATCTGCAATATGAGACCGAGAGAGGATCAATGAGCGGAGCTTCAAGCGGAGCTT 840

```


QY 849 TTCTGAGCCGAGGAGACCCCTGGG-ATGCTGACCTCTGCAAAAGGTCAATTTGGAGT 907
| | | | |
Db 841 TTCTGAGACGAGGAGAACCCCTGGAGATGCTGACCTTCTGAAAAGGTCCCATTTGGAGT 900
| | | | |
QY 908 CCAGCAGCAGACCCCA 923
| | | | |
Db 901 CCAGCAGCAGACCCCA 916
| | | | |
RESULT 5
BX400889 998 bp mRNA linear EST 13-MAY-2003
LOCUS BX400889 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DK003YL24 5-PRIME, mRNA sequence.
ACCESSION BX400889
VERSION BX400889.1 GI:30622359
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4968.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK003DF120P1&cluster=4968.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DK003DF120P1.
FEATURES
Source Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK003YL24"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 223 a 263 c 314 g 186 t 12 others
ORIGIN
Query Match 50.6%; Score 821; DB 13; Length 998;
Best Local Similarity 96.1%; Prid. No. 2.5e-164;
Matches 900; Conservative 4; Mismatches 26; Indels 7; Gaps 6;
QY 84 GAGAGCCGAGATGCTTGGCGGAGGCGAAGACTGGAGAGCCCGCGGAGAGCGCA 143
| | | | |
Db 57 GAGATGCCGAGATGCTTGGCGGAGGCGAAGACTGGAGAGCCCGCGGAGAGCGCA 115
| | | | |
QY 144 CCGTCCGACACGAGCGCCCTTCTGATGAGGGGTGAGCGGCGCACTGCCAGCGGAAATC 203
| | | | |
Db 116 CCGTCCGACACGAGCGCCCTTCTGATGAGGGGTGAGCGGCGCACTGCCAGCGGAAATC 175
| | | | |
QY 204 GACCGTGTGAGAAAGATCATGAGTGTCTGGACAGAGCAGGTGGAACAGCGGACGC 263
| | | | |
Db 176 GACCGTGTGAGAAAGATCATGAGTGTCTGGACAGAGCAGGTGGAACAGCGGACGC 235
| | | | |
QY 264 GAAAGTGTGATCTCTGAGCAGAGCAGGTTCTCAAGGTCTTGAACGCGCAGAGCAGAAAGC 323
| | | | |
Db 236 GAAAGTGTGATCTCTGAGCAGAGCAGGTTCTCAAGGTCTTGAACGCGCAGAGCAGAAAGC 295
| | | | |

QY 324 CAAGGCTTGAAGAGCAGTACAAATTTTGACCATCCAGATGCGCTTGATATGATTTGAT 383
| | | | |
Db 296 CAAGGCTTGAAGAGCAGTACAAATTTTGACCATCCAGATGCGCTTGATATGATTTGAT 355
| | | | |
QY 384 GCACAGACTCTGAAAGACATCGTGGAGGCGAAGACGCTGAGGTGCCAGCTATGATTT 443
| | | | |
Db 356 GCACAGACTCTGAAAGACATCGTGGAGGCGAAGACGCTGAGGTGCCAGCTATGATTT 415
| | | | |
QY 444 TGTGACACACTCAAGTTTACAGAGACCAACGCTGCTTACCTGCGGACGTGTTCTGTT 503
| | | | |
Db 416 TGTGACACACTCAAGTTTACAGAGACCAACGCTGCTTACCTGCGGACGTGTTCTGTT 475
| | | | |
QY 504 TGAAGGACATCTGTTGTTTCTTACAGCCAGAGAGATCCGGGACATGTTTCACTGGGCTCTT 563
| | | | |
Db 476 TGAAGGACATCTGTTGTTTCTTACAGCCAGAGAGATCCGGGACATGTTTCACTGGGCTCTT 535
| | | | |
QY 564 CCGTGAACACCGACTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623
| | | | |
Db 536 CCGTGAACACCGACTCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
| | | | |
QY 624 GAGGAGCTGGAGCAGATTTCTGACGACATCACACCTTCTGTAAGCCGCTTGGAGA 683
| | | | |
Db 596 GAGGAGCTGGAGCAGATTTCTGACGACATCACACCTTCTGTAAGCCGCTTGGAGA 654
| | | | |
QY 684 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743
| | | | |
Db 655 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
| | | | |
QY 744 GGTGCTCATCAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803
| | | | |
Db 715 GGTGCTCATCAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
| | | | |
QY 804 ATGGCACCGAGAGGCTCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
| | | | |
Db 775 ATGGCACCGAGAGGCTCCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
| | | | |
QY 864 CCAACCTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
| | | | |
Db 834 -CAACCTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
| | | | |
QY 924 CTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
| | | | |
Db 893 CTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949
| | | | |
QY 984 GCGTGGGAGCGCCACCAACACCACTGCTTCTCTC 1020
| | | | |
Db 950 AGCTGGGAGCGCCACCAACCACTGCTTCTCTC 986
| | | | |
RESULT 6
BG491358 860 bp mRNA linear EST 27-MAR-2001
LOCUS BG491358
DEFINITION 602535642F1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684738 5',
mRNA sequence.
ACCESSION BG491358
VERSION BG491358.1 GI:13452870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 860)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTP/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/JLNL at:
 http://image.jnl.gov
 Plate: L1CM1493 row: c column: 11
 High quality sequence stop: 844.
 Location/Qualifiers

FEATURES

Source

1. 860
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684738"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 195 a 223 c 288 g 154 t

ORIGIN

Query Match 49.0%; Score 795.2; DB 10; Length 860;
 Best Local Similarity 99.1%; Pred. No. 7.7e-159;
 Matches 852; Conservative 0; Mismatches 3; Indels 5; Gaps 5;

QY 44 CCGGGGAAAGGGGCGGCGGCGGAGACCCGATGCGCGGAGCGGAGCCGAGATGGCTTGG 103
 DB 2 CTGGGGAAGGGGCGGCGGCGGAGACCCGATGCGGAGCGGAGCCGAGATGGCTTGG 61
 QY 104 CGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCCGATGCGGAGCGGCGGCT 163
 DB 62 CGGAGGCGAAGACTGCGAGAGCCCGCGCGGAGCGGAGCCGATGCGGAGCGGCGGCT 121
 QY 164 TCCTGATAGGGGTAGCGCGGCGACCTGCCAGCGGAGATCGACCGTGTGTGAAGAATCA 223
 DB 122 TCCTGATAGGGGTAGCGCGGCGACCTGCCAGCGGAGATCGACCGTGTGTGAAGAATCA 181
 QY 224 TGGAGTTCCTGGGACAGAACGAGGTGTGAACGCGGAGCGGAGAGTGTCTCTGAGCC 283
 DB 182 TGGAGTTCCTGGGACAGAACGAGGTGTGAACGCGGAGCGGAGAGTGTCTCTGAGCC 241
 QY 284 AGGACAGGTTCTACAAAGTCTCTGACGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 343
 DB 242 AGGACAGGTTCTACAAAGTCTCTGACGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 301
 QY 344 ACAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGACA 403
 DB 302 ACAATTTTGAACCATCCAGATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGACA 361
 QY 404 TCGTGGAGGCGAAACGAGGTGTGAAGTGTGATTTGTGACACACTCAAGGTTAC 463
 DB 362 TCGTGGAGGCGAAACGAGGTGTGAAGTGTGATTTGTGACACACTCAAGGTTAC 421
 QY 464 CAGAGACCAAGGAGTGTACCCCTGCGGAGCGTGTCTGTTGAGGAGATCTTGTTTCT 523
 DB 422 CAGAGACCAAGGAGTGTACCCCTGCGGAGCGTGTCTGTTGAGGAGATCTTGTTTCT 481
 QY 524 ACAGCCAGAGAGATCCGAGACATGTTCCACTCTGCGCTTTGTTGAGACCGACTCCGAG 583
 DB 482 ACAGCCAGAGAGATCCGAGACATGTTCCACTCTGCGCTTTGTTGAGACCGACTCCGAG 541
 QY 584 TCAAGCTGTCTCTGAAG-AGTTCCTCGGAGCGTGTGCGGAGGAGGAGGAGGAGGAGGAGG 642
 DB 542 TCAAGCTGTCTCTGAAGAGTTCCTCGGAGCGTGTGCGGAGGAGGAGGAGGAGGAGGAGG 601
 QY 643 CTGACGAGTACACACCTTCTGTAAGCGGCGCTTTCGAGAGTCTGCTGTCGCGAGCAAG 702
 DB 602 CTGACGAGTACACACCTTCTGTAAGCGGCGCTTTCGAGAGTCTGCTGTCGCGAGCAAG 661
 QY 703 AAGTATCCGATGTGATCATCCACGAGAGGTGACAATA-TGGTTGCATCAACTGAT 761

DB 662 AAGTATCCGATGTGATCATCCACGAGAGGTGACAATAATTGTTGCCATCAACTGAT 721
 QY 762 CTGACGAGTACACACCTTCTGTAAGCGGCGCTTTCGAGAGTCTGCTGTCGCGAGCAAG 820
 DB 722 CTGACGAGTACACACCTTCTGTAAGCGGCGCTTTCGAGAGTCTGCTGTCGCGAGCAAG 781
 QY 821 CCATGAGGCGAGGAGTCAACGAGACCTTTTCTGAGCGAGGAGGAGGAGGAGGAGGAGGAGG 880
 DB 782 CCATGAGGCGAGGAGTCAACGAGACCTTTTCTGAGCGAGGAGGAGGAGGAGGAGGAGGAGG 840
 QY 881 CTTCT-GGCAACGCTCACA 899
 DB 841 CTTCTGGGCAACGCTCACA 860

RESULT 7
 BX394295
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 EST.
 BX394295.1 GI:30624219
 BX394295
 BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC013XE13 5-PRIME, mRNA sequence.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequenage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr; Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013AC07Q1&cluster=4968.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0DC013AC07Q1.

FEATURES

Source

1. 1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013XE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 288 a 284 c 363 g 226 t 40 others

ORIGIN

Query Match 48.8%; Score 792.4; DB 13; Length 1201;
 Best Local Similarity 93.4%; Pred. No. 3.2e-158;
 Matches 930; Conservative 7; Mismatches 40; Indels 19; Gaps 10;

QY 72 ATGCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 131
 DB 67 ATGCGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 125
 QY 132 GCCGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 191
 DB 126 GCCGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 184
 QY 192 CAGCGGAAAGTGCAGCGTGTGTGAAGAATCATGAGTTGCTTGGACAGAGAGGTGA 251

```

Db      185 CAGCGGGAATCGACCGTGTGTGAGAAATCATGAGTTGCTGGACAGAACGAGGTGA 244
Qy      252 ACAGCGGACCGGAAGGTGCTCATCTGAGCCAGAGACAGTTCTACAAAGTCTCTGAGCGC 311
Db      245 ACAGCGGACCGGAAGGTGCTCATCTGAGCCAGAGACAGTTCTACAAAGTCTCTGAGCGC 304
Qy      312 AGAGCAGAAAGCCCAAGGCTTTGAAAGGACAGTACATTTTGGACCATCCAGATGCTTTGA 371
Db      305 AGAGCAGAAAGCCCAAGGCTTTGAAAGGACAGTACATTTTGGACCATCCAGATGCTTTGA 364
Qy      372 TAAATGATTTGATGACAGAGACTCTGAAAGAAATCTGTGAGAGGCAAAACGCTGAGGTGCC 431
Db      365 TAAATGATTTGATGACAGAGACTCTGAAAGAAATCTGTGAGAGGCAAAACGCTGAGGTGCC 424
Qy      432 GACCTAATGATTTTGTGACACACTCAAGTTACAGAGCAACGCTGCTTACCTTGGGGA 491
Db      425 GACCTAATGATTTTGTGACACACTCAAGTTACAGAGCAACGCTGCTTACCTTGGGGA 484
Qy      492 CGTGGTCTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCA 551
Db      485 CGTGGTCTGTTTGAAGGACATCTTGTGTTCTACAGCCAGAGAGATCCGGGACATGTTCCA 544
Qy      552 CTTGCGGCTCTTCTGTGACACCGACCTCCAGCTCAGGCTGTCTTGAAGAGTTCTCCGGGA 611
Db      545 CTTGCGGCTCTTCTGTGACACCGACCTCCAGCTCAGGCTGTCTTGAAGAGTTCTCCGGGA 604
Qy      612 CTTGCGGCTCTTCTGTGACACCGACCTTGAAGAGTTCTTGAAGAGTTCTTGAAGAGG 671
Db      605 CTTGCGGCTCTTCTGTGACACCGACCTTGAAGAGTTCTTGAAGAGTTCTTGAAGAGG 664
Qy      672 GGCCTTGAAGAGATTTGCTGCTGCGACCAAGAAATATGCGAGTGTATCATCCAGAGG 731
Db      665 GGCCTTGAAGAGATTTGCTGCTGCGACCAAGAAATATGCGAGTGTATCATCCAGAGG 724
Qy      732 AGTGAACAATATGTTGCTGCATCAACTGATCTGTGACAGACATCCAGAGACTTGTGAATG 791
Db      725 AGTGAACAATATGTTGCTGCATCAACTGATCTGTGACAGACATCCAGAGACTTGTGAATG 784
Qy      792 TGACATCTGCAAAATGGAACCGAGAGGCTCCATATGGCGGAGCTTCAAGCGGACCTTTTC 851
Db      785 TGACATCTGCAAAATGGAACCGAGAGGCTCCATATGGCGGAG-CTACAAAGCGGACCTTTTC 843
Qy      852 TGAGCCAGGAGGACCACTGAGGATGCTGACCTCTGGGAAAACGCTCAATTTGAGGTCCAG 911
Db      844 TGAGCCAGGAGGAC-CACTCTGGAGTCTGACCTCTGGGAAAACGCTCAATTTGAGGT-CAAG 901
Qy      912 CAGCAGACCCCACTGAGGAGGCTGCCAGGCTTCAAGGAGGATCTTCCCGCGGACATGTGTG 971
Db      902 CAGCAGAC-CCCACTGAGGAGGCTGCC- -GAGCTCAGGAGGATCTCC- -GCCGGGACATGTGTN 955
Qy      972 TTTGAGGAGCTGAGGCTTGGGAGCGCCACCCACCACTGCTTCTTGGCGGACCCCA 1031
Db      956 TTTGAG- -GAGCTGAGGCTTGGGAGCGCCACCCACCACTGCTTCTTGGCGGACCCCA 1007
Qy      1032 GGGGAGTGTAGACGAGGAGGCTTCTTCACTCACTCAAG 1067
Db      1008 AGGGAGTGTAGAGGAGGCTTCTTMTTMAAGRTGGA 1043

```

```

RESULT 8
BM803129 1039 bp mRNA linear EST 05-MAR-2002
LOCUS     BM803129
DEFINITION AGENCOURT_6468837 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5721513
5', mRNA sequence.
ACCESSION BM803129
VERSION   BM803129.1 GI:19119952
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1039)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.

```

```

TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgabbs-remail.nih.gov
          Tissue Procurement: Invitrogen
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: LHAM12706 row: b column: 10
          High quality sequence stop: 652.
FEATURES
    source
        1..1039
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5721513"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_125"
        /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
        Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
        of three ovaries, from females ranging in age from 38 to
        49 yo. Library is oligo-dT primed and directionally cloned
        (each site is destroyed upon cloning). Average insert
        size 2.1 kb, insert size range 1-3.5 kb. Library is
        normalized and enriched for full-length clones and was
        constructed by C. Gruber (Invitrogen). Research Genetics
        tracking code 036."
BASE COUNT 240 a 266 c 309 g 220 t 4 others
ORIGIN
Query Match 48.3%; Score 783.6; DB 12; Length 1039;
Best Local Similarity 98.3%; Pred. No. 2.3e-156;
Matches 812; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

```

```

Qy      323 CCAAGGCTTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATTAAGATTGA 382
Db      1 CCAAGGCTTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATTAAGATTGA 60
Qy      383 TGACAGAGACTCTGAGAGACATGTTGAGGAGGCAAAACGCTGAGAGTGCAGCTTATGATT 442
Db      61 TGACAGAGACTCTGAGAGACATGTTGAGGAGGCAAAACGCTGAGAGTGCAGCTTATGATT 120
Qy      443 TTGTGACACACTCAAGTTACAGAGACCAAGGTGTTACCTTGGGACGTGTTCTGT 502
Db      121 TTGTGACACACTCAAGTTACAGAGACCAAGGTGTTACCTTGGGACGTGTTCTGT 180
Qy      503 TTGAGGAGCATCTTGTGTTCTTACAGCAGAGATCCGGGACATGTTCCAGCTGGCTCT 562
Db      181 TTGAGGAGCATCTTGTGTTCTTACAGCAGAGATCCGGGACATGTTCCAGCTGGCTCT 240
Qy      563 TCGTGAACACCGACTCCAGAGTCAAGGCTGTCTTGAAGAGTTCTCCGGGAGCTGCCGAG 622
Db      241 TCGTGAACACCGACTCCAGAGTCAAGGCTGTCTTGAAGAGTTCTCCGGGAGCTGCCGAG 300
Qy      623 GGAGGAGCTTGAAGAGATTTGACGCACTAACAACCTTGTGAAGCCGGCTTGAAG 682
Db      301 GGAGGAGCTTGAAGAGATTTGACGCACTAACAACCTTGTGAAGCCGGCTTGAAG 360
Qy      683 AGTTTCTGCTGCGGACCAAGAGATGATGATGATATCCCAAGAGAGTGAACAATA 742
Db      361 AGTTTCTGCTGCGGACCAAGAGATGATGATGATATCCCAAGAGAGTGAACAATA 420
Qy      743 TGGTTCATCAACTGATCTGTGACAGCATTCAGAGCAATTCGATGAGTGAACATCTGCA 802
Db      421 TGGTTCATCAACTGATCTGTGACAGCATTCAGAGCAATTCGATGAGTGAACATCTGCA 480
Qy      803 AATGACACGAGAGGCTCAATGAGCGGAGCTTAAAGCGGACCTTTTGTGAGCCAGGG 862
Db      481 AATGACACGAGAGGCTCAATGAGCGGAGCTTAAAGCGGACCTTTTGTGAGCCAGGG 540

```

QY 863 ACCACCTGGAGTGTGACTCTGCGAAACGGTACATTTGAGTCCAGCAGACACCCC 922
 Db 541 ACCACCTGGAGTGTGACTCTGCGAAACGGTACATTTGAGTCCAGCAGACACCCC 600
 QY 923 ACTGAGGGGCTGCGAGCTTCAGGAGAGTCTCCGCGCGAGATGTGTTTCAGGAGCTG 982
 Db 601 ACTAGGGGGCTGCGAGCTTCAGGAGAGTCTCCGCGCGAGATGTGTTTCAGGAGCTG 660
 QY 983 AGCCCTGGGAGCGCCACCCACCAACCACTGCTTCCCTCTGCGCGACCCCGAGGAGTGTTA 1042
 Db 661 AGCCTGAGGAGCGCCACCCACCAACCACTGCTTCCCTCTGCGCGACCCCGAGGAGTGTTA 720
 QY 1043 GCACGAGGCGCTTCTCTCAGAGAGTGG-AACTCAGATGTGTCATCAGACTCAACTT 1101
 Db 721 GCACGAGGCGCTTCTCTCAGAGAGTGGAAACTCAGATGTGTCATCAGACTCAACTT 780
 QY 1102 GCTGGGACATGACAGGCG--TTCTGAGGTTTTCAGGCACTTAAAG 1145
 Db 781 GCTGGGACATGACAGGCGGTTTCTGAGAGTTCAGGCACTTAAAG 826

RESULT 9
 LOCUS BG491384 869 bp mRNA linear EST 27-MAR-2001
 DEFINITION 602535670F1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684936 5',
 mRNA sequence.
 ACCESSION BG491384
 VERSION BG491384.1 GI:13452896
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 869)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1. 869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /class_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAACGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN

Query Match 48.1%; Score 781; DB 10; Length 869;
 Best Local Similarity 97.0%; Pred. No. 8,1e-156;
 Matches 837; Conservative 0; Mismatches 21; Indels 5; Gaps 4;
 QY 44 CCGGGAGAGGGCGGGCGCGGGAGCCGATGCGGGAGCGGAGCCGAGATGGCTTCGG 103

Db 2 CTGGGGAGAGGGCGGGCGCGGGAGCCGATGCCGAGAGCGAGGCCGAGATGGCTTCGG 61
 QY 104 CCGGAGGCGGAAGACTCGAGAGACCCCGCGCGGAGCCGATCCGCGACACCGGCTT 163
 Db 62 CCGGAGGCGGAAGACTCGAGAGACCCCGCGCGGAGCCGATCCGCGACACCGGCTT 121
 QY 164 TCCGTATGAGGGGTGAGCGGCGGCACTGCCAGCGGGAAGTGCACCGTGTGAAATCA 223
 Db 122 TCTGTATGAGGGGTGAGCGGCGGCACTGCCAGCGGGAAGTGCACCGTGTGAAATCA 181
 QY 224 TGAAGTTGCTGGGACAGAAAGAGGTGAAACAGCGGAGCGGAGGTTGTCATCTGAGCC 283
 Db 182 TGAAGTTGCTGGGACAGAAAGAGGTGAAACAGCGGAGCGGAGGTTGTCATCTGAGCC 241
 QY 284 AGGACAGGTTCTACAAAGTCTCTGACCGGACAGAGAAAGCCCTTGAAGACAT 343
 Db 242 AGGACAGGTTCTACAAAGTCTCTGACCGGACAGAGAAAGCCCTTGAAGACAT 301
 QY 344 ACAATTTTGACATCCAGATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGACA 403
 Db 302 ACAATTTTGACATCCAGATGCTTTGATTAATGATTTGATGACAGGACTCTGAAGACA 361
 QY 404 TCGTGAAGGCAAAACGGTGAAGTGCAGACTATGATTTTGTGACACACTCAAGGTTAC 463
 Db 362 TCGTGAAGGCAAAACGGTGAAGTGCAGACTATGATTTTGTGACACACTCAAGGTTAC 421
 QY 464 CAGAGACCAAGGTGTCTACCTCGCGGACGTGTTCTGTTTGAAGGCACTTGTGTTCT 523
 Db 422 CAGAGACCAAGGTGTCTACCTCGCGGACGTGTTCTGTTTGAAGGCACTTGTGTTCT 481
 QY 524 ACAGCCAGAGAGATCCGGAGATGTCACCTCGGCTCTTGTGAGACACCGACTCCGAG 583
 Db 482 ACAGCCAGAGAGATCCGGAGATGTCACCTCGGCTCTTGTGAGACACCGACTCCGAG 541
 QY 584 TCAAGCTGTCTGAAGAAGTCTTCGCGACGTGCGCGGAGGAGGACCTGAGACGATTC 643
 Db 542 TCAAGCTGTCTGAAGAAGTCTTCGCGACGTGCGCGGAGGAGGACCTGAGACGATTC 601
 QY 644 TGAGCGAGTACACCACTTCGCTGTAAGCCGCGCTTGTGAGAGTTCGCGGACCAAG- 702
 Db 602 TGAGCGAGTACACCACTTCGCTGTAAGCCGCGCTTGTGAGAGTTCGCGGACCAAG 661
 QY 703 AAGTATGCCGATGTGATCACTCCAGAGAGTGAACATAT-AGTTGCCATCAACTGAT 761
 Db 662 AAGTATGCCGATGTGATCACTCCAGAGAGTGAACATATAGTTGCCATCAACTGAT 721
 QY 762 CGTGACGACATCCAGACATTTGATGATGATCTGCAATGGCA-CCGAGAGGGT 820
 Db 722 CGTGACGACATCCAGACATTTGATGATGATCTGCAATGGCACCCGAGAGGGT 781
 QY 821 CCAATGGGGGAGACTCAAGCGGACCTTTTCTGAGCAGGGGACCACTCGGATGCTGA 880
 Db 782 CCAATGGGGGAGACTCAAGCGGACCTTTTCTGAGCAGGGG--ACCACTGGGATCTGA 839
 QY 881 CCTCTGGCAAAACGGTCAACTTTG 903
 Db 840 CCTCTGGCAAAACGGTCAACTTTG 862

RESULT 10
 LOCUS BG492079/c 929 bp mRNA linear EST 27-MAR-2001
 DEFINITION 602535670T1 NIH_MGC_41 Homo sapiens CDNA clone IMAGE:4684936 3',
 mRNA sequence.
 ACCESSION BG492079
 VERSION BG492079.1 GI:13453591
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 929)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L10M1493 row: k column: 17
 High quality sequence start: 24
 High quality sequence stop: 820.
 Location/Qualifiers

FEATURES

source

```

1..929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4684936"
/issue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 41"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGCGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

```

BASE COUNT 195 a 280 c 249 g 205 t

ORIGIN

Query Match 47.8%; Score 776.6; DB 10; Length 929;
 Best Local Similarity 96.0%; Pred. No. 7.1e-155;
 Matches 871; Conservative 0; Mismatches 29; Indels 7; Gaps 7;

QY 443 TTGTGACACACTCAAGGTTACAGAGACACGAGTGTCTACCTCGCGAGCTGTTCTGT 502
 DB 912 TTGGGAACACTCAAGTTACAGAGACCCGCGTGTCTACCTCGCACCGGGTGTGCT 853
 QY 503 TTGAGGCACTTGTGTCTGACG-GCAGAGATCCGGG-ACATGTTCCACCTGGCCCT 560
 DB 852 TCGAGGCAATTTGTGTCTGACGCGAGAGATCCGGGAACATGTTCCACCTGGCCCT 793
 QY 561 CTTCGTGG-ACACGACTCCGAC-GTCAGGCTGTCTGAAAG-AGTTCTCCGAGACGTG-C 616
 DB 792 CTTTGTGGAACCGGATCCGACCTGTCAAGCTGTCTGAAATGTTCTCCGGACGTGTC 733
 QY 617 GCCGAGGAGGAGCTTGAGAGCATTTCTGACGAGTACACCACTTGTGTGAAGCCGCT 676
 DB 732 GCCGAGGAGGAGCTTGAGAGCATTTCTGACGAGTACACCACTTGTGTGAAGCCGCT 673
 QY 677 TCGAGAGTTCTGTGCTGCGCAAGAAAGTATGCGATGTGATCATCCACGAGAGATTGG 736
 DB 672 TCGAGAGTTCTGTGCTGCGCAAGAAAGTATGCGATGTGATCATCCACGAGAGATTGG 613
 QY 737 ACAATATGTTGCTGATCAACCTGATCGTGAGCATTCAGAGCATTTCTGAATGTTGACA 796
 DB 612 ACAATATGTTGCTGATCAACCTGATCGTGAGCATTCAGAGCATTTCTGAATGTTGACA 553
 QY 797 TCTGCAATGAGCAGCAGAGAGGTTCAATGGGCGGAGCTACAGCGGACCTT-TTCTGAG 855
 DB 552 TCTGCAATGAGCAGCAGAGAGGTTCAATGGGCGGAGCTACAGCGGACCTTGTGTTGAG 493
 QY 856 CCGAGGAGACACCTTGAGAGTGTGACCTTGTGCGAAAGGTCACATTTGAGTCCAGAGC 915
 DB 492 CCGAGGAGACACCTTGAGAGTGTGACCTTGTGCGAAAGGTCACATTTGAGTCCAGAGC 433
 QY 916 AGACCCCACTGAGGAGGCTGCGAGACCTCAGGAGAGGTCTCCCGCCGAGCATGTGTCTCA 975

DB 432 AGACCCCACTGAGGAGGCTGCGAGACCTCAGGAGAGGTCTCCCGCCGAGCATGTGTCTCA 373
 QY 976 GGAATGAGCTTGGGAGACGGCCACACACCACTGCTTCTTGGCGGACCCAGGGG 1035
 DB 372 GGAATGAGCTTGGGAGACGGCCACACCACTGCTTCTTGGCGGACCCAGGGG 313
 QY 1036 AGTGTGAGCAGGAGCCTTCTCACTCAGAGATGGAATCTCAGATGTGCTCAGACT 1095
 DB 312 AGTGTGAGCAGGAGCCTTCTCACTCAGAGATGGAATCTCAGATGTGCTCAGACT 253
 QY 1096 CAATCTGCTGGAACACTGACAGGCGTTCTGAGTTTTCAGCCACTTGGCTGTTGCGG 1155
 DB 252 CAATCTGCTGGAACACTGACAGGCGTTCTGAGTTTTCAGCCACTTGGCTGTTGCGG 193
 QY 1156 TTTAAAGATCCCTCTAGTCTCAGTGAATGCGAAGATGTCAGAGACCTGGAGGC 1215
 DB 192 TTTAAAGATCCCTCTAGTCTCAGTGAATGCGAAGATGTCAGAGACCTGGAGGC 133
 QY 1216 TTTCTGAGAGATGTGAGGACATTTATTTGGGAAATTTGAGAGACAGCTGACACTGCG 1275
 DB 132 TTTCTGAGAGATGTGAGGACATTTATTTGGGAAATTTGAGAGACAGCTGACACTGCG 73
 QY 1276 TGACCTGATGTTTGTGTTGACAGTGAACCCACAGTGGAGAGAGTGTTCAGTGTGATC 1335
 DB 72 TGACCTGATGTTTGTGTTGACAGTGAACCCACAGTGGAGAGAGTGTTCATATACCGCGC 13
 QY 1336 TGCTTCT 1342
 DB 12 GGGTCT 6

RESULT 11

BO072501

LOCUS 1036 bp mRNA linear EST 02-APR-2002
 DEFINITION AGENCOURT 6838909 NIH_MGC_122 Homo sapiens cDNA IMAGE:5761694
 ACCESSION BO072501
 VERSION BO072501.1 GI:19901547
 KEYWORDS EST.

SOURCE ORGANISM

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1036)

NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L10M12810 row: 1 column: 15
 High quality sequence stop: 644.
 Location/Qualifiers

FEATURES

source

```

1..1036
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5761694"
/lab_host="DH10B"
/clone_id="NIH MGC 122"
/notes="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for

```


Db 360 ATGCCAGATGTCAGAGAGCCCTGGAGGCTTCTGTGAGATGTGAGCATTATT 419
 Qy 1244 GGGGAATTAGAGAGAGCTTAGACA CTGGCTGGCCGTGATGTTTGTGACAGTGAAC 1303
 Db 420 GGGGAATTAGAGAGAGCTTAGACA CTGGCTGGCCGTGATGTTTGTGACAGTGAAC 479
 Qy 1304 CACAGTGGAGAGAGATTTTTCAGTCTGATCTGTTCTTACACACTCACA CATTAAT 1363
 Db 480 CACAGTGGAGAGAGATTTTTCAGTCTGATCTGTTCTTACACACTCACA CATTAAT 539
 Qy 1364 CAAAAGTTTGTGAACAAGTACTTCTTTTATCATGTACATGCTCATGTTTCTG 1423
 Db 540 CAAAAGTTTGTGAACAAGTACTTCTTTTATCATGTACATGCTCATGTTTCTG 599
 Qy 1424 TTTTCTGTTTATACACAGAGCTGTGTGGCTTACAAACCTTAATTTTCAATGCTG 1483
 Db 600 TTTTCTGTTTATACACAGAGCTGTGTGGCTTACAAACCTTAATTTTCAATGCTG 659
 Qy 1484 GTTTCAGTCCAGGCTGGCTTACACAGATATGGGAGCCACTGAGGATGTTTCCCCC 1543
 Db 660 GTTTCAGTCCAGGCTGGCTTACACAGATATGGGAGCCACTGAGGATGTTTCCCCC 719
 Qy 1544 TTGCTGTGCTTAAAGCAGAGAGCGAGCGGATGCCCTGAGACACCGCATCACA 1603
 Db 720 CTGCTTGTGCTTAAAGCAGAGAGCGAGCGGATGCCCTGAGACACCGCATCACA 779
 Qy 1604 CCCAGGCTTGTGCGGCGCCAG 1624
 Db 780 CCCAGGCTTGTGCGGCGCCAG 800

RESULT 13

LOCUS B0719741 956 bp mRNA linear EST 16-JUL-2002
 DEFINITION A58NCOURT_8229663 Lupski_dorsal_root_ganglion Homo sapiens cDNA
 B0719741
 accession IMAGE:6184625 5', mRNA sequence.

B0719741.1 GI:21858638
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 956)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.jnl.gov
 Plate: LNA13574 row: b column: 18
 High quality sequence stop: 552.

FEATURES

source
 1. 956
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6184625"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /note="Vector: PCMV-SpOxT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACGCGTCCG-3' and

BASE COUNT

202 a 272 c 296 g 186 t

ORIGIN

Query Match 46.6%; Score 756.4; DB 13; Length 956;
 Best Local Similarity 95.9%; Pred. No. 1.4e-150;
 Matches 798; Conservative 0; Mismatches 31; Indels 3; Gaps 2;

5'-GACTAGTCTAGATCCGAGCGCCCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

88 GCCGAGATGGCTTGGCGGGAGGCGAAGCTGGAGAGGCCCGCGCGGAGCGCCGT 147
 Db 1 GCCGAGATGGCTTGGCGGGAGGCGAAGCTGGAGAGGCCCGCGCGGAGCGCCGT 60
 Qy 148 CCGCACCAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGAAC 207
 Db 61 CCGCACCAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAGTGAAC 120
 Qy 208 GTGTGTGAAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAAGCGGCAAG 267
 Db 121 GTGTGTGAAGATCATGAGTGTCTGGGACAGACGAGGTGGAACAAGCGGCAAG 180
 Qy 268 GTGTGTATCTGATCCAGAGACAGATCTCAAGAGTCTGACCGGACAGAGAGCGCAAG 327
 Db 181 GTGTGTATCTGATCCAGAGACAGATCTCAAGAGTCTGACCGGACAGAGAGCGCAAG 240
 Qy 328 GCCTTGAAGAGACATGATATTTTGAACATCCAGATGCTTGTATATGATTTGATGAC 387
 Db 241 GCCTTGAAGAGACATGATATTTTGAACATCCAGATGCTTGTATATGATTTGATGAC 300
 Qy 388 AGGACTCTGAAGAACTCTGTGAGGCGAAGACGCTGAGGTGCGGACCTATGATTTTGTG 447
 Db 301 AGGACTCTGAAGAACTCTGTGAGGCGAAGACGCTGAGGTGCGGACCTATGATTTTGTG 360
 Qy 448 ACACACTCAAGTTACAGAGACACAGGTGTATCCCTGCGGACGATGCTTGTGTGAG 507
 Db 361 ACACACTCAAGTTACAGAGACACAGGTGTATCCCTGCGGACGATGCTTGTGTGAG 420
 Qy 508 GGCATCTGTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTTCTGTG 567
 Db 421 GGCATCTGTGTGTTCTACAGCCAGAGATCCGGGACATGTTCCAGCTGCGCTTCTGTG 480
 Qy 568 GACACCGACTCCGAGCTCAGGCTGTCTGAAAGTTCTCCGAGACGTGCGGAGGAGG 627
 Db 481 GACACCGACTCCGAGCTCAGGCTGTCTGAAAGTTCTCCGAGACGTGCGGAGGAGG 540
 Qy 628 GACCTGAGACATTTGTAGACGAGTACACACCTTGTGAAAGCGGCGCTTCCAGAGTTTC 687
 Db 541 GACCTGAGACATTTGTAGACGAGTACACACCTTGTGAAAGCGGCGCTTCCAGAGTTTC 600
 Qy 688 TGCCTGCGGACAAAGAGATGTCGAGTGTATCATCCACAGAGAGTGAACAATATAGG-T 746
 Db 601 TGCCTGCGGACAAAGAGATGTCGAGTGTATCATCCCGAGAGATGGAACAATATAGGTT 660
 Qy 747 TGCATCAACCTGATGTCGAGACATCCAGACATTTGGAATGTGACATTCGAATG 806
 Db 661 TGCATCAACCTGATGTCGAGACATCCAGACATTTGGAATGTGACATTCGAATG 720
 Qy 807 GCACCGAGAGAGTCCAAATGGGGGAGCTACAAAGCGGACCTTTTCTGAGCC--AGGAGC 864
 Db 721 GCACCGAGAGAGTCCAAATGGGGGAGCTTCAAAGGGGACCTTTTCTTAACCCAGGGGAAAC 780
 Qy 865 CACCTGGAGATGCTGCTCTGGCAAAAGGTGACATTTGAGATCCAGACGA 916
 Db 781 CCCCCTGGATGTTGACCTTTGGCCCAAGTTCCATTTTGGAGGGCCAGCA 832

RESULT 14
 LOCUS B0459108 1051 bp mRNA linear EST 05-FEB-2002

```

DEFINITION  AGENCOURT_6414543_NIH_MGC_72_Homo_sapiens_cDNA_clone IMAGE:5557433
5', mRNA sequence.
ACCESSION   BM459108
VERSION     BM459108.1 GI:18508148
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1051)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DRP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LLM12278 row: m column: 18
            High quality sequence stop: 542.
            Location/Qualifiers
                1..1051
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5557433"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 72"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
BASE COUNT  237 a      264 c      312 g      221 t      17 others
ORIGIN
Query Match      46.5%; Score 755.2; DB 12; Length 1051;
Best Local Similarity 97.5%; Pred. No. 2.6e-150;
Matches 788; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

QY      89 CCGAGATGCGCTCCGCGGGAAGGACCTCGAGAGCCCGCGGAGGCGGACCGC 148
Db      1 CCGAGATGCGCTCCGCGGGAAGGACCTCGAGAGCCCGCGGAGGCGGACCGC 60

QY      149 CGCACACAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCG 208
Db      61 CGCACACAGCGGCCCTTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAGTCGACCG 120

QY      209 TGTGTGGAAGATATGATGAGTTTGTCTGGACAGAAAGAGTGGAAACAGCGGAGCGGAAAGG 268
Db      121 TGTGTGGAAGATATGATGAGTTTGTCTGGACAGAAAGAGTGGAAACAGCGGAGCGGAAAGG 180

QY      269 TGGTCATCTGAGCGCAGGACAGTCTTCTACAGGTCCTGACGCGAGAGAGGAGGCGCAAGG 328
Db      181 TGGTCATCTGAGCGCAGGACAGTCTTCTACAGGTCCTGACGCGAGAGAGGAGGCGCAAGG 240

QY      329 CTTTGAAGAAGACAGTACATTTTGAACATCCAGATGCTTGTGATGATTTGATGACACA 388
Db      241 TCTTGAAGAAGACAGTACATTTTGAACATCCAGATGCTTGTGATGATTTGATGACACA 300

QY      389 GGAATCTGGAAGAATCGTGGAGGGCAAAACGGTGGAGGTGCCAGCTATGATTTTGTGA 448
Db      301 GGAATCTGGAAGAATCGTGGAGGGCAAAACGGTGGAGGTGCCAGCTATGATTTTGTGA 360

QY      449 CACACTCAAGGTTACAGAGACACAGGTCGTCTACCCCTGCGGAGCGTGGTTCTGTTTGAAG 508
Db      361 CACACTCAAGGTTACAGAGACACAGGTCGTCTACCCCTGCGGAGCGTGGTTCTGTTTGAAG 420

QY      509 GCATCTTGTGTGTTCTACAGCGAGAGATCGGGAATGTTTCCACTGCGCTTTCGTGG 568

```

```

Db      421 GCATCTTGTGTGTTCTACAGCGAGAGATCGGGAATGTTTCCACTGCGCTTTCGTGG 480
QY      569 ACAACCGATCCGACAGTCAGGCTGTCTCGAAGATTTCTCCGGGACGTGCCGAGGAGG 628
Db      481 ACAACCGATCCGACAGTCAGGCTGTCTCGAAGATTTCTCCGGGACGTGCCGAGGAGG 540
QY      629 ACCTGAGCAGATTTGACACCACTTCTGTAAGCGGCTTGAAGATTC 688
Db      541 ACCTGAGCAGATTTGACACCACTTCTGTAAGCGGCTTGAAGATTC 600
QY      689 GCCTGCCGACAAAGAGTATCCGATGTATCATCCACGAGAGTGGACATATGCTTG 748
Db      601 GCTGCCGACAAAGAGTATGATGTATCATCCACGAGAGTGGACATATGCTTG 660
QY      749 CATTCAACCTGATGTGACACATCATCAGAGCACTTCTGAATGTTGATCTGCAATGCG 808
Db      661 ACATCAACCTGATGTGACACATCATCAGAGCACTTCTGAATGTTGATCTGCAATGCT 720
QY      809 ACCGAGAGAGGTCCTAATGGCGGAGCTACAG- CGACCTTTCTGAGCCAGGAGCA- 866
Db      721 ACCGAGAGAGGTCCTAATGGCGGAGCTACAGAGTCCGACCTTTCTGATGAGGAGGAC 780
QY      867 CCTGGGATCTGACCTTGGCAACGG 894
Db      781 CCTGGGATCTGACCTTGGCAACGG 808

RESULT 15
CD514811
LOCUS
DEFINITION AGENCOURT_14376196_NIH_MGC_181_Homo_sapiens_cDNA_clone
IMAGE:30356201 5', mRNA sequence.
CD514811
VERSION     CD514811.1 GI:31446529
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 898)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bidg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: NDMA477 row: 1 column: 10
            High quality sequence stop: 662.
            Location/Qualifiers
                1..898
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:30356201"
                /tissue_type="White Matter"
                /dev_stage="Unknown"
                /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                /clone_lib="NIH MGC 181"
                /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
                (destroyed): Library is oligo-dT primed and directionally
                cloned (EcoRV site is destroyed upon cloning). Average
                insert size 1.42 kb. Library was constructed by
                (Invitrogen). Note: this is a NIH_MGC Library."
FEATURES
Source

```


BASE COUNT 194 a 231 c 310 g 158 t 5 others
ORIGIN

Query Match 46.4%; Score 753.4; DB 14; Length 898;
Best Local Similarity 99.9%; Pred. No. 6.1e-150;
Matches 754; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 23 CGCTGGGCGGGCCGCGCGGGGAGGGGGCGGGGAGCCCGATGCGCGGAG 82
Db 1 CGCTGGGCGGGCCGCGCGGGGAGGGGGCGGGGAGCCCGATGCGCGGAG 60
QY 83 CGGAGGCCGAGATGCTTCCGCGGAGCGAAGACTGCGAGGCCCGCGGAGCCG 142
Db 61 CGGAGGCCGAGATGCTTCCGCGGAGCGAAGACTGCGAGGCCCGCGGAGCCG 120
QY 143 ACCGTCGCGACCGAGCGGCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAT 202
Db 121 ACCGTCGCGACCGAGCGGCTTCTCTGATAGGGGTGAGCGCGGCACTGCCAGCGGAAAT 180
QY 203 CGACCGGTGTGAAGAATCATGGAGTTGCTGGGACAGAAAGAGTGAACAGCGGAGC 262
Db 181 CGACCGGTGTGAAGAATCATGGAGTTGCTGGGACAGAAAGAGTGAACAGCGGAGC 240
QY 263 GGAAGGTGTATCTTGAGCCAGGACAGTCTTACAAGTCTGACCGGACAGCAGAGG 322
Db 241 GGAAGGTGTATCTTGAGCCAGGACAGTCTTACAAGTCTGACCGGACAGCAGAGG 300
QY 323 CCAAGGCTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGA 382
Db 301 CCAAGGCTTGAAAGGACAGTACATTTTGAACCATCCAGATGCTTTGATATGATTTGA 360
QY 383 TGACACGAGACTGTGAAGACATGTGAGGGCAAAAGGTGAGGTGCCGACCTATGAT 442
Db 361 TGACACGAGACTGTGAAGACATGTGAGGGCAAAAGGTGAGGTGCCGACCTATGAT 420
QY 443 TTGTGACACACTGAAGTTTACAAGACACAGGTGTCTACCTGCGGACGTGTTCTGT 502
Db 421 TTGTGACACACTGAAGTTTACAAGACACAGGTGTCTACCTGCGGACGTGTTCTGT 480
QY 503 TTGAGGGCATCTTGTTCTTAACAGCCAGAGATCCGGAATGTTCCACTGCGCTCT 562
Db 481 TTGAGGGCATCTTGTTCTTAACAGCCAGAGATCCGGAATGTTCCACTGCGCTCT 540
QY 563 TCGTGAACAACGACTCCGAGTCAAGGCTGTGGAAGTTTCCGGGACGTGCCGAG 622
Db 541 TCGTGAACAACGACTCCGAGTCAAGGCTGTGGAAGTTTCCGGGACGTGCCGAG 600
QY 623 GGAAGGACTGAGCAGATTCTGACGAGTACACCACTTCGTGAAGCCGGCTTCGAGG 682
Db 601 GGAAGGACTGAGCAGATTCTGACGAGTACACCACTTCGTGAAGCCGGCTTCGAGG 660
QY 683 AGTTCTGCTGCGACCAAGAAGTATGCCATGTATCATCCACGAGAGTGAACATA 742
Db 661 AGTTCTGCTGCGACCAAGAAGTATGCCATGTATCATCCACGAGAGTGAACATA 720
QY 743 TGGTTGCAATCAACTGATGATGAGACATCCAG 777
Db 721 TGGTTGCAATCAACTGATGATGAGACATCCAG 755
```

Search completed: November 25, 2003, 02:00:45
Job time : 3919.04 sec

This Page Blank (uspto)

Db 310 TCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAGAAACATGTTGAGGGCAA 369
QY 417 AAGGTGAGAGTGGCCGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACACGGT 476
Db 370 AAGGTGAGAGTGGCCGACCTATGATTTTGTGACACACTCAAGTTTACAGAGACACGGT 429
QY 477 GGCTACCCCTGGAGAGTGGTCTGTTTGGAGGATCTTGATGTTTCTACAGCAGGAGAT 536
Db 430 GGCTACCCCTGGAGAGTGGTCTGTTTGGAGGATCTTGATGTTTCTACAGCAGGAGAT 489
QY 537 CCGGAGCATGTTCCACCTGCGCTCTTCTGTGACACCGACTCCGACGTGAGCTGTCTCG 596
Db 490 CCGGAGCATGTTCCACCTGCGCTCTTCTGTGACACCGACTCCGAGGTGAGCTGTCTCG 549
QY 597 AAGAGTTCTCCGGAGAGTGGTGGCCGAGGAGGACCTGAGACAGATTTTGAACGACATCAC 656
Db 550 AAGAGTTCTCCGGAGAGTGGTGGCCGAGGAGGACCTGAGACAGATTTTGAACGACATCAC 609
QY 657 CACCTTCGTGAGACCGGCTTCGAGAGAGTCTGCTGCTCGACAAAGAGTATGCCGATGT 716
Db 610 CACCTTCGTGAGACCGGCTTCGAGAGAGTCTGCTGCTCGACAAAGAGTATGCCGATGT 669
QY 717 GATCATTCACGAGAGAGTGGACATATGTTGCCATCACTGATCTGTGAGCAGACATCCA 776
Db 670 GATCATTCACGAGAGAGTGGACATATGTTGCCATCACTGATCTGTGAGCAGACATCCA 729
QY 777 GGACATT 783
Db 730 GGACATT 736

RESULT 3
LOCUS BM917506 1044 bp mRNA linear EST 12-MAY-2002
DEFINITION AGENCOURT_6606568 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5484033
5', mRNA sequence.
BM917506
BM917506.1 GI:19367885
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1044)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Place: LNCM2012 row: k column: 10
High quality sequence stop: 624.
Location/Qualifiers
1. 1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5484033"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/note="Organ: blood. Vector: pORF7. Site 1: XhoI. Site 2:
EcoRI. cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

FEATURES

source

1. 1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5484033"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/note="Organ: blood. Vector: pORF7. Site 1: XhoI. Site 2:
EcoRI. cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."
BASE COUNT 224 a 306 c 318 g 196 t
ORIGIN

Query Match 43.6%; Score 708; DB 12; Length 1044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAGAGCGAAAGACTGCGAGAGAGCCCGCGCGAGAGCCGACCGTCCGACCAAGCGCCCTTC 165
Db 1 GAGAGCGAAAGACTGCGAGAGAGCCCGCGCGAGAGCCGACCGTCCGACCAAGCGCCCTTC 60
QY 166 CTGATGAGGGGTGAGCGCGCGGACCTGCGAGCGGGAAGTCCACCGTGTGTGAAGAAATCATG 225
Db 61 CTGATGAGGGGTGAGCGCGCGGACCTGCGAGCGGGAAGTCCACCGTGTGTGAAGAAATCATG 120
QY 226 GAGTTGCTGGGACAGAAAGAGGTGGAACAGCGGCGAGAGAGGTGTCATCTCTGAGCCAG 285
Db 121 GAGTTGCTGGGACAGAAAGAGGTGGAACAGCGGCGAGAGAGGTGTCATCTCTGAGCCAG 180
QY 286 GACAGGTTCTACAGAGTCTGACAGGACAGAGCCAGAGCCCTTGAAGGACATGAC 345
Db 181 GACAGGTTCTACAGAGTCTGACAGGACAGAGCCAGAGCCCTTGAAGGACATGAC 240
QY 346 AATTTTGACATCCAGATGCTCTTGTATATGATTTTGTGACACAGATCTGAAAGAAATC 405
Db 241 AATTTTGACATCCAGATGCTCTTGTATATGATTTTGTGACACAGATCTGAAAGAAATC 300
QY 406 GTGAGAGGCAAAACGAGTGGAGTGGCCGACCTATGATTTTGTGACACATCAAGGTTACCA 465
Db 301 GTGAGAGGCAAAACGAGTGGAGTGGCCGACCTATGATTTTGTGACACATCAAGGTTACCA 360
QY 466 GAGACCAAGGTGATCTTACCTGCGGACGTGTTCTGTTTGAAGGCACTTGTGTTCTAC 525
Db 361 GAGACCAAGGTGATCTTACCTGCGGACGTGTTCTGTTTGAAGGCACTTGTGTTCTAC 420
QY 526 AGCAGAGAAATCCGGGACATGTTCCACCTGCGGCTCTTGTGAGACACCGACCTCCAG 585
Db 421 AGCAGAGAAATCCGGGACATGTTCCACCTGCGGCTCTTGTGAGACACCGACCTCCAG 480
QY 586 AGGCTGTCTGAAAGAGTCTCCGGGACGTGCGCCCGAGGAGGAACTGTGAGCAGATTCTG 645
Db 481 AGGCTGTCTGAAAGAGTCTCCGGGACGTGCGCCCGAGGAGGAACTGTGAGCAGATTCTG 540
QY 646 AGCAGATACACCACTCTGTGAAGCGGCTTGAAGAGTCTGCTGCGACAAAGAG 705
Db 541 AGCAGATACACCACTCTGTGAAGCGGCTTGAAGAGTCTGCTGCGACAAAGAG 600
QY 706 TATGCCGATGATATCATCCAGAGAGTGGACATATGTTGCCATCAACTGATCTG 765
Db 601 TATGCCGATGATATCATCCAGAGAGTGGACATATGTTGCCATCAACTGATCTG 660
QY 766 CAGCAGATCAGAGACATCTGAATGAGTGCATCTGCAATGACACCGA 813
Db 661 CAGCAGATCAGAGACATCTGAATGAGTGCATCTGCAATGACACCGA 708

FEATURES

source

RESULT 4
LOCUS BX372390/c 941 bp mRNA linear EST 08-MAY-2003
DEFINITION BX372390 Homo sapiens B CELLS (RAMOS CELL LINE) COR 25-NORMALIZED
Homo sapiens cDNA clone CSDBL009YD15 3-PRIME, mRNA sequence.
ACCESSION BX372390
VERSION BX372390.1 GI:30456111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization


```

QY 226 GAGTTCGTGGACAGAAAGAGTGGAAACAGCGGACGGAGAGTGTCTATCTGAGCCAG 285
    |||
Db 212 GAGTTGCTGGACAGAAAGAGTGGAAACAGCGGACGGAGAGTGTCTATCTGAGCCAG 271
    |||
QY 286 GACAGGTTCTACAAAGTCTTGAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
    |||
Db 272 GACAGGTTCTACAAAGTCTTGAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 331
    |||
QY 346 AATTTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGAACATC 405
    |||
Db 332 AATTTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGAACATC 391
    |||
QY 406 GTGAGAGGCAAAACGGTGGAGAGTGGACCGACCTATGATTTTGTGACACACTGAAGTTTCCA 465
    |||
Db 392 GTGAGAGGCAAAACGGTGGAGAGTGGACCGACCTATGATTTTGTGACACACTGAAGTTTCCA 451
    |||
QY 466 GAGACCAACGAGTGTCTACCCCTGCGGACGTGTCTGTTTGAAGGACATCTTGATTTCTAC 525
    |||
Db 452 GAGACCAACGAGTGTCTACCCCTGCGGACGTGTCTGTTTGAAGGACATCTTGATTTCTAC 511
    |||
QY 526 AGCCAGAGAGATCCGGGACATGTTTCCACTGCGCTCTTCTGTGACACCCGACTCCGAGCTC 585
    |||
Db 512 AGCCAGAGAGATCCGGGACATGTTTCCACTGCGCTCTTCTGTGACACCCGACTCCGAGCTC 571
    |||
QY 586 AGGCTGTCTGAGAGAGTCTCCGGGACGTGCGGACCGAGAGAGAGAGAGAGAGAGAGAGAG 645
    |||
Db 572 AGGCTGTCTGAGAGAGTCTCCGGGACGTGCGGACCGAGAGAGAGAGAGAGAGAGAGAGAG 631
    |||
QY 646 AGCCAGTACACACCTTCTGTGAAAGCCGCTTGAAGAGAGTCTGCTGCGGACAAAGAG 705
    |||
Db 632 AGCCAGTACACACCTTCTGTGAAAGCCGCTTGAAGAGAGTCTGCTGCGGACAAAGAG 691
    |||
QY 706 TATGCCGATGTATCATCCACGAGAGAGTGGACATAT 743
    |||
Db 692 TATGCCGATGTATCATCCACGAGAGAGTGGACATAT 729
    |||

```

```

RESULT 6
LOCUS BG770518 730 bp mRNA linear EST 15-MAY-2001
DEFINITION 60274236F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4859577 5',
    mRNA sequence.
ACCESSION BG770518
VERSION BG770518.1 GI:14081171
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLES Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: csabbs-remail.nih.gov
          Tissue Procurement: ATCC/DCPD/DRP
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LHCMI715 row: h column: 10
          High quality sequence stop: 720.
          Location/Qualifiers
            1..730

```

```

FEATURES
    source
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4859577"
        /tissue_type="melanotic melanoma, high MDR (cell line)"
        /lab_host="DH10B (phage-resistant)"
        /clone_1ib="NIH_MGC_49"

```

```

/Note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the Laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

```

BASE COUNT 156 a 190 c 258 g 126 t
ORIGIN
Query Match 42.4% Score 688; DB 12; Length 720;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 GGGGAAAGGGCGGGCGGGGAGACCCGATGCGCGGAGCGGAGCCGAGATGGCTTGAGCG 105
    |||
Db 34 GGGGAAAGGGCGGGCGGGGAGACCCGATGCGCGGAGCGGAGCCGAGATGGCTTGAGCG 93
    |||
QY 106 GGAAGGGAAGACTGCGAGAGACCCCGCGCGGAGCCGACCCGTCGACACAGCGGCGCTTC 165
    |||
Db 94 GGAAGGGAAGACTGCGAGAGACCCCGCGCGGAGCCGACCCGTCGACACAGCGGCGCTTC 153
    |||
QY 166 CTGATAGGGGTGAGCGCGGCGACCTGCGACCGGGAAGTCGACCGTGTGTGAGGAATCATG 225
    |||
Db 154 CTGATAGGGGTGAGCGCGGCGACCTGCGACCGGGAAGTCGACCGTGTGTGAGGAATCATG 213
    |||
QY 226 GAGTTCCTGGACAGAACAGGTGGAAACAGCGGACCGGAGAGTGTCTATCTGAGCCAG 285
    |||
Db 214 GAGTTCCTGGACAGAACAGGTGGAAACAGCGGACCGGAGAGTGTCTATCTGAGCCAG 273
    |||
QY 286 GACAGGTTCTACAAAGTCTTGAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
    |||
Db 274 GACAGGTTCTACAAAGTCTTGAACGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 333
    |||
QY 346 AATTTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGAACATC 405
    |||
Db 334 AATTTTGACCATCCAGATGCTTTGATATGATTTGATGACAGAGACTCTGAAGAACATC 393
    |||
QY 406 GTGAGAGGCAAAACGGTGGAGAGTGGACCGACCTATGATTTTGTGACACACTGAAGTTTCCA 465
    |||
Db 394 GTGAGAGGCAAAACGGTGGAGAGTGGACCGACCTATGATTTTGTGACACACTGAAGTTTCCA 453
    |||
QY 466 GAGACCAACGAGTGTCTACCCCTGCGGACGTGTCTGTTTGAAGGACATCTTGATTTCTAC 525
    |||
Db 454 GAGACCAACGAGTGTCTACCCCTGCGGACGTGTCTGTTTGAAGGACATCTTGATTTCTAC 513
    |||
QY 526 AGCCAGAGAGATCCGGGACATGTTTCCACTGCGCTTGTGAGAGAGAGAGAGAGAGAGAGAG 585
    |||
Db 514 AGCCAGAGAGATCCGGGACATGTTTCCACTGCGCTTGTGAGAGAGAGAGAGAGAGAGAGAG 573
    |||
QY 586 AGGCTGTCTGAGAGAGTCTCCGGGACGTGCGGACCGAGAGAGAGAGAGAGAGAGAGAGAG 645
    |||
Db 574 AGGCTGTCTGAGAGAGTCTCCGGGACGTGCGGACCGAGAGAGAGAGAGAGAGAGAGAGAG 633
    |||
QY 646 AGCCAGTACACACCTTCTGTGAAAGCCGCTTGAAGAGAGTCTGCTGCGGACAAAGAG 705
    |||
Db 634 AGCCAGTACACACCTTCTGTGAAAGCCGCTTGAAGAGAGTCTGCTGCGGACAAAGAG 693
    |||
QY 706 TATGCCGATGTATCATCCACGAGAGAG 733
    |||
Db 694 TATGCCGATGTATCATCCACGAGAGAG 721
    |||

```

```

RESULT 7
LOCUS B1256928 725 bp mRNA linear EST 17-JUL-2001
DEFINITION 60297553F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114816 5',
    mRNA sequence.
ACCESSION B1256928
VERSION B1256928.1 GI:14811813
KEYWORDS EST.

```

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 725)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM11279 row: c column: 09
 High quality sequence stop: 707.
 Location/Qualifiers
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5114816"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies."
 BASE COUNT 175 a 181 c 224 g 145 t
 ORIGIN
 Query Match 41.7%; Score 677; DB 12; Length 725;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 197 GGAAGTGCACCGCTGTGAGAAGATCATGTGAGTCTGGGACAGAACAGGTGGAACAGC 256
 Db 1 GGAAGTGCACCGCTGTGAGAAGATCATGTGAGTCTGGGACAGAACAGGTGGAACAGC 60
 257 GGCAGCGGAAGGTGTCATCTGAGCCAGAGCAGGTTCTACAAGTCTCTGACGACAGC 316
 Db 61 GGCAGCGGAAGGTGTCATCTGAGCCAGAGCAGGTTCTACAAGTCTCTGACGACAGC 120
 317 AGAAGGCCAGGCTTTGAAAGACAGTCAATTTTACCATCCAGATGCTTTGATATG 376
 Db 121 AGAAGGCCAGGCTTTGAAAGACAGTCAATTTTACCATCCAGATGCTTTGATATG 180
 377 ATTGATGACAGACCTGTAAGAACATCGTGAAGGGGCAAAACGTTGAGGTGCCGACT 436
 Db 181 ATTGATGACAGACCTGTAAGAACATCGTGAAGGGGCAAAACGTTGAGGTGCCGACT 240
 437 ATGATTTTGTGACACTCAAGTTTACAGAGACCAAGCGTGTCTACCTTCGCGAGCTGG 496
 Db 241 ATGATTTTGTGACACTCAAGTTTACAGAGACCAAGCGTGTCTACCTTCGCGAGCTGG 300
 497 TTCTGTTTGAAGGCACTTGTGTCTTACAGCCAGAGATCGGGAACATGTTCCACTGC 556
 Db 301 TTCTGTTTGAAGGCACTTGTGTCTTACAGCCAGAGATCGGGAACATGTTCCACTGC 360
 557 GCCTCTTGTGAGACACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGC 616
 Db 361 GCCTCTTGTGAGACACCGACTCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTGC 420
 617 GCCAGAGGAGGACCTTGAGAGATTCTGACGACGATACACCACTTGTGAGAGCGGACT 676
 Db 421 GCCAGAGGAGGACCTTGAGAGATTCTGACGACGATACACCACTTGTGAGAGCGGACT 480
 677 TCGAGAGTTTGCCTGCGCAAAAGAAATATGCGATGTGATCATCCAGAGAGTGG 736

Db 481 TCGAGAGTTTCTGCTGCTCCGACAAAGATATGCCATGTGATCATCCAGAGAGTGG 540
 737 ACATATGTTTGCATTAACCTGATGCTGAGACATCCAGACATTTCTTAATGTGACA 796
 Db 541 ACATATGTTTGCATTAACCTGATGCTGAGACATCCAGACATTTCTTAATGTGACA 600
 797 TCTGCAATATGACACCGAGAGGTTCCAAATGGGCGAGCTCAAGCGGACCTTTCTGAGC 856
 Db 601 TCTGCAATATGACACCGAGAGGTTCCAAATGGGCGAGCTCAAGCGGACCTTTCTGAGC 660
 857 CAGGGGACCACTGGG 873
 Db 661 CAGGGGACCACTGGG 677
 RESULT 8
 BG491384 869 bp mRNA linear EST 27-MAR-2001
 LOCUS 60253567F1 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:4684936 5',
 DEFINITION mRNA sequence.
 ACCESSION BG491384.1 GI:13452896
 VERSION BG491384
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 869)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLM1493 row: k column: 17
 High quality sequence stop: 863.
 Location/Qualifiers
 1..869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4684936"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_41"
 /note="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 197 a 223 c 291 g 156 t 2 others
 ORIGIN
 Query Match 40.6%; Score 659; DB 10; Length 869;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 46 GGGGAAGGGCGGGCGGCGGACCCGATGCGCGGAGCGGAGCCGAGATGCTTCGGCG 105
 Db 4 GGGGAAGGGCGGGCGGCGGAGCCGATGCGCGGAGCGGAGCCGAGATGCTTCGGCG 63
 106 GAGGCGAAGACTGCGAAGACCCCGCGCGGAGGCGGACCGTCCGACCAAGCGGCCCTTC 165
 Db 64 GAGGCGAAGACTGCGAAGACCCCGCGCGGAGGCGGACCGTCCGACCAAGCGGCCCTTC 123


```

QY 166 CTGATAGGGGGAGCGGGGCACTGGCAGCGGGGAAGTCSA CCGTGTGTGAGGAAGATCATG 225
Db 124 CTGATAGGGGGAGCGGGGCACTGGCAGCGGGGAAGTCSA CCGTGTGTGAGGAAGATCATG 183
QY 226 GAGTGTCTGGAGACGAA CGAGGTGAA CAGCGCGGAGGAGGTGATCTTGAGGCAG 285
Db 184 GAGTGTCTGGAGACGAA CGAGGTGAA CAGCGCGGAGGAGGTGATCTTGAGGCAG 243
QY 286 GACAGGTTCTACAGGTCCTGACCGGACAGACAGAAAGCCAAAGCCTTGAAAGACAGTAC 345
Db 244 GACAGGTTCTACAGGTCCTGACCGGACAGACAGAAAGCCAAAGCCTTGAAAGACAGTAC 303
QY 346 AATTTGACCATCAGATGCTTTGATATGATTTGATGACAGGATCTGGAAGAACATC 405
Db 304 AATTTGACCATCAGATGCTTTGATATGATTTGATGACAGGATCTGGAAGAACATC 363
QY 406 GTGAGGGGCAAAA CGGTGAGGTGCGACCTATGATTTGTGACACACTCAAGTTACCA 465
Db 364 GTGAGGGGCAAAA CGGTGAGGTGCGACCTATGATTTGTGACACACTCAAGTTACCA 423
QY 466 GAGACCA CGGTGCTTACCTCGCGGACGTGTTCTGTTGAGGACATCTTGTTCTAC 525
Db 424 GAGACCA CGGTGCTTACCTCGCGGACGTGTTCTGTTGAGGACATCTTGTTCTAC 483
QY 526 AGCAGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGACACCGACCTCCGACGTC 585
Db 484 AGCAGAGATCCGGGACATGTTCCACTGCGCTCTTCGTGACACCGACCTCCGACGTC 543
QY 586 AGGCTGTCTGGAAGATTTCTCCGGGACGTGCGCGGAGGAGGACCTGAGACAGATTCTG 645
Db 544 AGGCTGTCTGGAAGATTTCTCCGGGACGTGCGCGGAGGAGGACCTGAGACAGATTCTG 603
QY 646 ACCGAGTACACCACTTGTGAAAGCCGGCCTTGAGGAGTTCTGCTGCCGACAAAGAA 704
Db 604 ACCGAGTACACCACTTGTGAAAGCCGGCCTTGAGGAGTTCTGCTGCCGACAAAGAA 662

```

```

RESULT 9
LOCUS BU838101 883 bp mRNA linear EST 16-OCT-2002
DEFINITION AGENCOURT 8071586 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088251
BU838101
ACCESSION BU838101 GI:24022496
VERSION BU838101.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 883)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LNCM2326 row: C column: 04
High quality sequence stop: 655.
Location/Qualifiers
1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088251"
/risue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

FEATURES

source

```

1..883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6088251"
/risue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

```

/clone_11b="NIH MGC 112"
/notes="Organ: skin; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library"
BASE COUNT 188 a 240 c 272 g 183 t
ORIGIN

```

```

Query Match 40.2%; Score 653; DB 13; Length 883;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 753; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 457 AGGTTACAGAGACCAACGGTGTCTACCTCGCGGACGTGTTCTGTTGAGGACATCTTG 516
Db 2 AGGTTACAGAGACCAACGGTGTCTACCTCGCGGACGTGTTCTGTTGAGGACATCTTG 61
QY 517 GTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTCGCTCTTCTGTGACACCGAC 576
Db 62 GTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTCGCTCTTCTGTGACACCGAC 121
QY 577 TCCGACGTACAGGCTGTCTGGAAGATTTCCGGGACGTGCGCGGAGGAGGACCTGGAG 636
Db 122 TCCGACGTACAGGCTGTCTGGAAGATTTCCGGGACGTGCGCGGAGGAGGACCTGGAG 181
QY 637 CAGATTCTGACGACGATACCAACCTTGTGAAAGCCGGCCTTGAGGAGTTCTGCTGCGG 696
Db 182 CAGATTCTGACGACGATACCAACCTTGTGAAAGCCGGCCTTGAGGAGTTCTGCTGCGG 241
QY 697 ACAAAGATGATGCCGATGTGATCATCCACGAGAGTGAACAATGCTTGGTCATCAAC 756
Db 242 ACAAAGATGATGCCGATGTGATCATCCGAGAGTGAACAATGCTTGGTCATCAAC 301
QY 757 CTGATGTGACGACATCCAGGACATTTGAAATGAGACATCTGCAAAATGGACCGAGGA 816
Db 302 CTGATGTGACGACATCCAGGACATTTGAAATGAGACATCTGCAAAATGGACCGAGGA 361
QY 817 GGGTCCAAATGGGCGGAGCTTACCAAGCCGACCTTCTGAGCCAGGGGACCACTTGGATG 876
Db 362 GGGTCCAAATGGGCGGAGCTTACCAAGCCGACCTTCTGAGCCAGGGGACCACTTGGATG 421
QY 877 CTGACCTCTGCGCAAAAGCTTCACTTTGAGTCCAGACGACACCCACTGAGGGGCTGCC 936
Db 422 CTGACCTCTGCGCAAAAGCTTCACTTTGAGTCCAGACGACACCCACTGAGGGGCTGCC 481
QY 937 GAGCCTCAGGGGACGTCCTCCGCGGACATGTGTTCAGGGGACAGGCTGGGGAGCGCC 996
Db 482 GAGCCTCAGGGGACGTCCTCCGCGGACATGTGTTCAGGGGACAGGCTGGGGAGCGCC 541
QY 997 CACCCACACCACTGCTTCTCTCGCGGACACCCAGGGGAGTGTAGCAGCAGGCTTTC 1056
Db 542 CACCCACACCACTGCTTCTCTCGCGGACACCCAGGGGAGTGTAGCAGCAGGCTTTC 601
QY 1057 CTGACTCAGAGGTGAAACTCAGATGTGACTCAGACTCAACTTCTGGGACACTGACA 1116
Db 602 CTGACTCAGAGGTGAAACTCAGATGTGACTCAGACTCAACTTCTGGGACACTGACA 661
QY 1117 GGGTTCCTGAGGTTTTCAGCACTTGAAGGCTGGTGAAGATCCCTCTAGTCA 1176
Db 662 GGGTTCCTGAGGTTTTCAGCACTTGAAGGCTGGTGAAGATCCCTCTAGTCA 721
QY 1177 CTGAGAAATGCCACAGAAATGTGAGGAGACCTGGG 1211
Db 722 CTGAGAAATGCCACAGAAATGTGAGGAGACCTGGG 756

```

```

RESULT 10
LOCUS B0072501 1036 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6838909 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5761694

```

5', mRNA sequence.
 ACCESSION BC072501
 VERSION BC072501.1 GI:19901547
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1036)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished
 JOURNAL Contact: Robert Strassberg, Ph.D.
 COMMENT Email: c9apbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
 Plate: LHAM12810 row: 1 column: 15
 High quality sequence stop: 644.
 Location/Qualifiers
 1..1036
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5761694"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_122"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SportE; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung; 16 week female spleen and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MGC Library."
 BASE COUNT 221 a 285 c 346 g 184 t
 ORIGIN
 Query Match 40.0%; Score 650; DB 13; Length 1036;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 650; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 46 GGGGAAAGGGCGGGCGGGGACCGGATGCGGGAGCGAGCCGAGATGGCTTCGCGG 105
 DB 16 GGGGAAAGGGCGGGCGGGGACCGGATGCGGGAGCGAGCCGAGATGGCTTCGCGG 75
 QY 106 GGAGGCGAAGCTCGAGAGCCCGCGCGGAGCCGACCTCCGACACAGCGGCCCTTC 165
 DB 76 GGAAGCGAAGCTCGAGAGCCCGCGCGGAGCCGACCTCCGACACAGCGGCCCTTC 135
 QY 166 CTGATAGGGGTGAGCGCGGCACTGCCAGCGGAGTTCGACCGTGTGTGAGAGATCATG 225
 DB 136 CTGATAGGGGTGAGCGCGGCACTGCCAGCGGAGTTCGACCGTGTGTGAGAGATCATG 195
 QY 226 GAGTTGCTGGAGCAAGACGAGTGAACAGCGGACCGGAGGAGTGTATCTTGAGGCAG 285
 DB 196 GAGTTGCTGGAGCAAGACGAGTGAACAGCGGACCGGAGGAGTGTATCTTGAGGCAG 255
 QY 286 GACAGGTTCTACAGGTCCTGACCGGAGAGAGCAAGGCGCTTGAAGAGCACTAC 345
 DB 256 GACAGGTTCTACAGGTCCTGACCGGAGAGCAAGGCGCTTGAAGAGCACTAC 315
 QY 346 AATTTTACCATCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAAGACATC 405
 DB 316 AATTTTACCATCCAGATGCTTTGATATGATTTGATGACAGAGCTTGAAGACATC 375
 QY 406 GTGAGGCGAAAAAGGTGAGAGTGCAGACTATGATTTGTGACACACTCAAGGTTACCA 465

|||||
 DB 376 GTGAGGCGAAAAAGGTGAGAGTGCAGACTATGATTTGTGACACACTCAAGTTACCA 435
 QY 466 GAGACCAAGGTGATCAACCTTCGGAGACGTGCTTGTGAGAGGACATCTTGTTCTAC 525
 DB 436 GAGACCAAGGTGATCAACCTTCGGAGACGTGCTTGTGAGAGGACATCTTGTTCTAC 495
 QY 526 AGCAGAGAGATCCGGGACATGTTCCACCTTCGCTTCGTGACACCCGACCTCGAGTC 585
 DB 496 AGCAGAGAGATCCGGGACATGTTCCACCTTCGCTTCGTGACACCCGACCTCGAGTC 555
 QY 586 AGGCTGTTCTGAAGAGTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGCATTTCTG 645
 DB 556 AGGCTGTTCTGAAGAGTCTCCGGGACGTGCGCCGAGAGGAGGACCTGAGCATTTCTG 615
 QY 646 AGCAGAGACCAACCTTCGAGACCGGCGCTTCGAGAGAGTTCGCTGCC 695
 DB 616 AGCAGAGACCAACCTTCGAGAGCGGCGCTTCGAGAGAGTTCGCTGCC 665
 RESULT 11
 EX394295
 LOCUS 1201 bp mRNA linear EST 13-MAY-2003
 DEFINITION BX394295 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 ACCESSION BX394295
 VERSION BX394295.1 GI:30624219
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS I.J.W.B., Gruber,C., Jesse,J. and Polyes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to the sequence cluster 4968.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC013AC070P1&cluster=4968.r Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC013AC070P1.
 Location/Qualifiers
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE13"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 288 a 284 c 363 g 226 t 40 others
 ORIGIN
 Query Match 39.6%; Score 643; DB 13; Length 1201;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 693; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 139 GCCGACCGTCCGACCAAGGCGCCCTTCGATAGGGGTGAGCGCGCACTGCGACGGG 198
 DB 133 GCCGACCGTCCGACCAAGGCGCCCTTCGATAGGGGTGAGCGCGCACTGCGACGGG 191
 QY 199 AAGTCACCGTGTGTGAAGATGATGAGTGTCTGGACAGAACGAGGTGAAACGCGG 258

```

Db      192 AAGTCAGCCGTGTGTGAGAAATCATGAGTTGCTGGAGAGAAAGAGTGTGAACACGCG 251
Qy      259 CACGGAGAGTGTGATCCTGAGCCAGAGAGAGTCTTCAAGAGTCTTGAACGGAGAGCAG 318
Db      252 CAGGNAAGGTGTGATCTGAGCCAGAGAGGTTCTCAAGAGTCTTGAACGGAGAGCAG 311
Qy      319 AAGGCCAAGGCTTGAAGAAGACAGTAAATTTTGAACCATCCAGATGCTTTGATATGAT 378
Db      312 AAGGCCAAGGCTTGAAGAAGACAGTAAATTTTGAACCATCCAGATGCTTTGATATGAT 371
Qy      379 TTGATGACAGAGACTCTGAAGAATCATGTGAGAGGCAAAAACGGTGAAGTCCGACCTAT 438
Db      372 TTGATGACAGAGACTCTGAAGAATCATGTGAGAGGCAAAAACGGTGAAGTCCGACCTAT 431
Qy      439 GATTTTGTACACACTAAAGTTAACAGAGAACCCAGGTGTCTTACCTGGGAGAGTGTGTT 498
Db      432 GATTTTGTACACACTAAAGTTAACAGAGAACCCAGGTGTCTTACCTGGGAGAGTGTGTT 491
Qy      499 CTGTTTGAAGGACATCTTGTTGTCTTACAGCCAGAGAGATCCGGACATGTTCCACCTGCG 558
Db      492 CTGTTTGAAGGACATCTTGTTGTCTTACAGCCAGAGAGATCCGGACATGTTCCACCTGCG 551
Qy      559 CTCTTGTGAGACACCGACTCCGACGTGAGGCTGTCTGAAAGATTCTCCGGAGAGTGC 618
Db      552 CTCTTGTGAGACACCGACTCCGACGTGAGGCTGTCTGAAAGATTCTCCGGAGAGTGC 611
Qy      619 CGAGGAGGAGACCTTGAGAGATTTCTGACGACGACACCTTCTGTAAGCCGCGCTTC 678
Db      612 CGAGGAGGAGACCTTGAGAGATTTCTGACGACGACACCTTCTGTAAGCCGCGCTTC 671
Qy      679 GAGAGATTCTGACCTGCGACAAAGATATGCGATGATCATCCACGAGAGAGTGGAC 738
Db      672 GAGAGATTCTGACCTGCGACAAAGATATGCGATGATCATCCACGAGAGAGTGGAC 731
Qy      739 AATATGTTTCCATCAACTGATGTGTCAGACATCCAGAGACATTTGATGATGATC 798
Db      732 AATATGTTTCCATCAACTGATGTGTCAGACATCCAGAGACATTTGATGATGATC 791
Qy      799 TGCATAATGACCCGAGAGAGGTCCAAATGGGCGA 832
Db      792 TGCATAATGACCCGAGAGAGGTCCAAATGGGCGA 825

```

```

RESULT 12
BX349263      894 bp      mRNA      linear      EST 05-MAY-2003
LOCUS      BX349263 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL009YD15 5-PRIME, mRNA sequence.
ACCESSION      BX349263
VERSION      BX349263.1 GI:30375361
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 894)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 4968.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAG057ZAI1 CS05457.1&cluster=4968.r.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/InvitrogenCorporation1600
            Fairday Avenue Genoscope sequence ID : CS0BAG057ZAI1_CS05457_1.
            location/Qualifiers
            1..894

```

FEATURES

source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YD15"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      192 a      242 c      253 g      206 t      1 others
ORIGIN

```

```

Query Match      39.0%; Score 634; DB 13; Length 894;
Best Local Similarity 99.9%; Pred. No. 8, 9e-311;
Matches 684; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      859 GGGGACACCCCTGGAGAGTGTGACCTCTGGCAAAACGGTCAATTGGAGTCCAGACAGAG 918
Db      35 GGGGACACCCCTGGAGAGTGTGACCTCTGGCAAAACGGTCAATTGGAGTCCAGACAGAG 94
Qy      919 CCCACTGAGGGGCTCCGAGCCTCAGGGCAGGCTCCCGCCGCGATGTGTTCAGGG 978
Db      95 CCCACTGAGGGGCTCCGAGCCTCAGGGCAGGCTCCCGCCGCGATGTGTTCAGGG 154
Qy      979 ACTGAGCTGGGAGCGCCCAACCCACCACTGCTCTCTGTGGCGCACCCAGGGAGT 1038
Db      155 ACTGAGCTGGGAGCGCCCAACCCACCACTGCTCTCTGTGGCGCACCCAGGGAGT 214
Qy      1039 GTTAGAGGAGGCGCTTCTCTCACTGAGAGTGGAACTAGATGTCTCACTGAGTCAA 1098
Db      215 GTTAGAGGAGGCGCTTCTCTCACTGAGAGTGGAACTAGATGTCTCACTGAGTCAA 274
Qy      1099 CTTCCTGGGACACTGACAGGCGTTCTCTGAGGTTTTCAGACATTTAGCTGTTCGGTTT 1158
Db      275 CTTCCTGGGACACTGACAGGCGTTCTCTGAGGTTTTCAGACATTTAGCTGTTCGGTTT 334
Qy      1159 AAGATCCCTCTTAGTGTCACTGAGAAATGCCACAAATGTGACAGAGAGCTGGAGGCTTC 1218
Db      335 AAGATCCCTCTTAGTGTCACTGAGAAATGCCACAAATGTGACAGAGAGCTGGAGGCTTC 394
Qy      1219 TGTGAGAAATGTGAGAGCACTATTGAGGAAATGTGAGAGACAGCCTTACACTGTGCTGG 1278
Db      395 TGTGAGAAATGTGAGAGCACTATTGAGGAAATGTGAGAGACAGCCTTACACTGTGCTGG 454
Qy      1279 CCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGG 1338
Db      455 CCTGATGTTTGTGACAGTGAACCCACAGTGGAGAGAGTTTTCAGTCTGATCTGG 514
Qy      1339 TTCTTACACACTGACACACATTAATGCAAAAGTTTGTGAACAAGTCTTTTCTTTTAC 1398
Db      515 TTCTTACACACTGACACACATTAATGCAAAAGTTTGTGAACAAGTCTTTTCTTTTAC 574
Qy      1399 AAGTTCACATGCTCTCAATGTTTCTGTTTCTGTTTATATACCAAGGCTGTGGGCT 1458
Db      575 AAGTTCACATGCTCTCAATGTTTCTGTTTCTGTTTATATACCAAGGCTGTGGGCT 634
Qy      1459 ACAACCTAATTTTCATGACCCAGTGTGTCAGTCCAGGCTGACCTTACAGATATGGGG 1518
Db      635 ACAACCTAATTTTCATGACCCAGTGTGTCAGTGTGTCAGGCTGACCTTACAGATATGGGG 694
Qy      1519 AGCCACTGAGGAGATGTTTCCCGCC 1543
Db      695 AGCCACTGAGGAGATGTTTCCCGCC 719

```

```

RESULT 13
BM545603      1055 bp      mRNA      linear      EST 20-FEB-2002
LOCUS      BM545603
DEFINITION      AGENCOURT 6500279 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5587864
            5', mRNA sequence.

```

ACCESSION	BM545603
VERSION	BM545603.1
KEYWORDS	GI:18777841
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
JOURNAL	1 (bases 1 to 1055)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished
	Contact: Robert Strausberg, Ph.D.
	Email: cgapbs-remail.nih.gov
	Tissue Procurement: Invitrogen
	cDNA Library Preparation: Life Technologies, Inc.
	Genomic Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
	DNA Sequencing by: Agencourt Bioscience Corporation
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LNLN at:
	http://image.llnl.gov
	Plate: L1AM12358 row: a column: 17
FEATURES	High quality sequence stop: 579.
source	1..1055
	Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5587864"
/lab_host="DH10B"
/clone_id="NTH_MGC_125"
/notice="Organ: ovary (pool of 3); Vector: PCMV-SPORT6;
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
Size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

Query Match	38.8%	Score 630	DB 12	Length 1055
Best Local Similarity	100.0%	Pred. No. 1e-308		
Matches 630	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	46	GGGGAAGGGCGGGCGCCGGGAACTCCGATGCCGCGGAGGCGGAGGCCGAGATGGCTTCGGCG	105	
Db	1	GGGGAAGGGCGGGCGCGGGGACCCGAGTCCGCGGAGGCGGAGGCGGAGATGGCTTCGGCG	60	
QY	106	GGAGGCGAAGACTCGAGAGCCCGCGCGCGGAGGCGGACCCGTCGCAACAGCGGCCCTTC	165	
Db	61	GGAGGCGAAGACTCGAGAGCCCGCGCGCGGAGGCGGACCGTCCGCAACAGCGGCCCTTC	120	
QY	166	CTGATTAGGGGTGACCGCGTGGCACTGCCAGCGGGAAGTGCACCGTGTGTGAAGATCATG	225	
Db	121	CTGATTAGGGGTGACCGCGTGGCACTGCCAGCGGGAAGTGCACCGTGTGTGAAGATCATG	180	
QY	226	GAGTTGCTGGGACAGAAAGAGTGGAAACAGCGGACGCGGAGAGTGGTATCTTGAGCCAG	285	
Db	181	GAGTTGCTGGGACAGAAAGAGTGGAAACAGCGGACGCGGAGAGTGGTATCTTGAGCCAG	240	
QY	286	GACAGGTTCTTCAAGGCTCTTGACGGCAGACAGAAAGGCCCAAGGCGCTTGAAGAACAAGTAC	345	
Db	241	GACAGGTTCTTCAAGGCTCTTGACGGCAGACAGAAAGGCCCAAGGCGCTTGAAGAACAAGTAC	300	
QY	346	AATTTTGACATCCAGATGCTTTTGATATATATTTGATGACAGGACTCTGAACAATC	405	
Db	301	AATTTTGACATCCAGATGCTTTTGATATATATTTGATGACAGGACTCTGAACAATC	360	
QY	406	GTGAGGCGCAAAACGGGTGAGAGGTGCCACCTATGATTTTGTGAACAACATAAGGTTACCA	465	
Db	361	GTGAGGCGCAAAACGGGTGAGAGGTGCCACCTATGATTTTGTGAACAACATAAGGTTACCA	420	

QY	466	GAGACCAACGGAGGCTCTACCCCGCGGACGAGGATTCGTTAGAGGGACATCTGGTCTTAC	525
Db	421	GAGACCAACGGAGGCTCTACCCCGCGGACGAGGATTCGTTAGAGGGACATCTGGTCTTAC	480
QY	526	AGCCAGAGAGATCCGGGACATGTTCCACTGCGCCTCTTCGTGACACCCAGATCCGACGTC	585
Db	481	AGCCAGAGAGATCCGGGACATGTTCCACTGCGCCTCTTCGTGACACCCAGATCCGACGTC	540
QY	586	AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGAGGACCTTGAGACATTTCTG	645
Db	541	AGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGAGAGGACCTTGAGACATTTCTG	600
QY	646	ACGCAGTACACACCTTCGTGAAGCCGGCC	675
Db	601	ACGCAGTACACACCTTCGTGAAGCCGGCC	630

RESULT	14
AU131406	
LOCUS	AU131406
DEFINITION	779 bp mRNA linear EST 01-AUG-2007
ACCESSION	AU131406 NT2RP3 Homo sapiens cDNA clone NT2RP3002519 5', mRNA sequence.
VERSION	AU131406
KEYWORDS	AU131406.1 GI:10991760
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE
AUTHORS
Oca, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Euharyota, Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 779)

TITLE HRI human CDNA project
JOURNAL Unpublished
COMMENT Contact: Takao Isogai

Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES	SOURCE	Location/Qualifiers
	1. .779	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="NT2RP3002519"	
	/cell_type="retinocarcinoma"	
	/cell_line="NT2"	
	/clone_idb="NT2RP3"	
	/note=Vector: pME18SFU3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"	
BASE COUNT	179 a 197 c 253 g 147 t 3 others	
ORIGIN		

Query Match	38.7%	Score 628	DB 9	Length 779
Best Local Similarity	99.9%	Pred. No. 9.6e-308		
Matches 678	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	71	GATGCGCGGAGCGGAGCGGAGATGCTTCGCGGAGGCGAGACTGCGAGACCCCG	130	
Db	1	GATGCGCGGAGCGGAGCGGAGATGCTTCGCGGAGGCGAGACTGCGAGACCCCG	60	
QY	131	CGCGGAGGCGGACCGTCCGCACTCAGCGGCGCTTCTCGATGAGGGGTGAGCGCGGCACTG	190	
Db	61	CGCGGAGGCGGACCGTCCGCACTCAGCGGCGCTTCTCGATGAGGGGTGAGCGCGGCACTG	120	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: November 25, 2003, 07:39:21 / Search time 16.8478 Seconds

(without alignments)
9066.011 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 536
Sequence: 1 gtcggggtcgccgcaccc.....ccagctgtgtcgggccag 1624

Scoring table:

	OLIGO
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 127863 seqs, 47026705 residues

Word size: 15

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame.n2p.model -DEV=xlh
-Q/cg2.1/USPTO.spool/US09896522/runat_21112003_184144_3222/app_query.faeta_1.2830
-DB=SwissProt 41 -QMT=fastan -SUFFIX=n2p_oli.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL
-OUTPM=pro -NORM=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.@CG2.1.1.33@runat_21112003_184144_3222 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -XONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query	Match length	ID	Description
1	277	51.7	277	1 UCK1_HUMAN	09h47 homo sapien
2	45	8.4	277	1 UCK1_MOUSE	p5263 mus musculu
3	25	4.7	261	1 UCK1_HUMAN	09bzx2 homo sapien
4	25	4.7	261	1 UCK2_MOUSE	09p93 mus musculu
5	15	2.8	260	1 UCK_DROME	09vc99 drosophila

ALIGNMENTS

RESULT 1
ID UCK1_HUMAN STANDARD; PRT; 277 AA.
AC 09h47;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine monophosphokinase 1) (Cytidine monophosphokinase 1).

GN UCK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Nord A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
human uridine-cytidine kinases.";
RN Mol. Pharmacol. 59:1181-1186 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Ho Y.S., Johnson R.K.;
RT "Human uridine kinase from prostate cancer cell line (LNCap).";
RN Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RA Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RN Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xin Y.R., Yu L., Zhao S.Y.;
RT "Cloning of a new human cDNA similar to Mus musculus uridine kinase
cDNA.";
RN Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: Phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate. Does not phosphorylate
deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
as a phosphate donor. Can also phosphorylate cytidine and uridine
nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
benzoylcytidine, 5-fluorocytidine, 2-chlorocytidine, 5-
methylcytidine, and N(4)-anisoylcytidine.
CC - CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC - CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC - PATHWAY: Pyrimidine salvage pathway.
CC - TISSUE SPECIFICITY: Ubiquitous.
CC - SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF237290; AAK28324.1; -
DR EMBL; AF254133; AAK49122.1; -
DR EMBL; AK022317; BAB14010.1; -
DR EMBL; AF125106; AAL75943.1; -
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PRO0988; URIDINKINASE.
DR TIGRFA; TIGR00235; udk; 1.
KW Transferase; Kinase; ATP-binding.
FT NP BIND 30 37
FT CONFLICT 8 17
FT CONFLICT 56 57
FT CONFLICT 247 247
SQ SEQUENCE 277 AA; 31434 MW; AFD9ED92780CD502 CRC64;
DCESAPPEND -> GARRAGAN (IN REF. 4).
QR -> HG (IN REF. 4).
SR -> T (IN REF. 4).

Alignment Scores:

Pred. No.:	1e-272	Length:	277
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.68%	Indels:	0
DB:	1	Gaps:	0

US-09-896-522-1 (1-1624) x UCK1_HUMAN (1-277)

```

QY 94 ATGGCTTCGGGGAGGAGGCGAAGCTGCGAGAGCCCGGCGGAGGCGGACCGTCCGCAC 153
Db 1 MetalaaserlagllygllygluaspCygegluSerProalaprProgluAlaasprgProHis 20
QY 154 CAGGGGCCCCCTTCGATAGGGGGTGAAGCGGCGGCACTGCGACGGGGAAGTGCACCTGTGT 213
Db 21 GlnmrgrProheleuilegllyvaliserGlYglYthrAlaaserGlYlySerThrValCys 40
QY 214 GAGAAAGATCATGAGATTGCTGGAGCAGAACGAGGTGGAACAGCGGCGGAGGAGTGC 273
Db 41 GlulyslleuettGlulLeuenglYGlmsnsluValGluglnmrgrGlnmrgrlyvalVal 60
QY 274 ATCCGAGCCAGGACAGGTTCTACAAAGTCCCTGACGCGAGACAGAAAGCCAGCCCTTG 333
Db 61 lleuSerGlmaasprgphetylrlvalleuthrlaglInlYslalylalaleu 80
QY 334 AAAAGACATGACATTTTGGACATCCAGATCCCTTGTATATGATTTGATGACAGAGACT 393
Db 81 LysglYglntYrAanpheasprhAsproasprAlaheasprAanAsprleuMechlAsrgThr 100
QY 394 CTGAGAAACATCGTGGAGGCGCAAAACGGTGGAGGTGCGGACCTTGTATTTGTGACACAC 453
Db 101 LeuylasentllevalgluInlylYethrValgluValProthrYrAsprhevalThHis 120
QY 454 TCAAGTTACAGAGACACCGTGGTCTACCCCTCGGACGTGTCTGTGTTGAGGGCATC 513
Db 121 SerArgleuProgluInlYrThrValValYrProAlaAsprValleuPhegluInlyle 140
QY 514 TTGGTGTCTACACCCAGGAGATCCGGGACATGTTCCACCTGGCCCTCTTGTGTGAGACAC 573
Db 141 LeuvalPleYrYserGlnglnlAlarAsprMetPhehlslleuAryleuPhevalAsprThr 160
QY 574 GACTCCGACGTGAGGCTGTCTCGAAGAGTTCTCCGGGACGTGCGCCGAGGGAGGAGCATC 633
Db 161 AsperAsprValaaglyleuSerArgrYValleuArAsprValaaglylYArAsprleu 180
QY 634 GAGCAGATTCTGACGACGATACACCACTTCGTGAAGCCGCGCTTCGAGAGATTGCTGCTG 693
Db 181 GluInlilleuThrInlYrThrThrPheVallyspProAlaPhegluInlYpHeCysleu 200
QY 694 CCGCAAAAGAGTATGCCGATGTATCATCCACGAGAGAGTGAACATATGTGTCATC 753
Db 201 ProthrYrYsYrYrAlaAsprVallelleProArgrglYValaAsprAsprMevalAlaile 220
QY 754 AACCTGATCGGAGCAGCATCCAGACATTTGATGAGTGTGATCATCTGCAAAATGCGACCGA 813
Db 221 AsnleuilevalglInhlslleglInAsprlleuAsnnglyAsprlleCylYslYstrPhlAsrg 240
QY 814 GAGAGGTCGAATGGCGGAGGAGTCAACAGCGAACCTTTTCTGAGCGAGGGAGCACCTTGCG 873
Db 241 GlyylYserAanglyArYsSerYrlyArgrhPheSerGlupProglYAsprhlAsprGlY 260
QY 874 ATGCTGACCTTGGCAAAACGTCACATTTGAGATGCCAGACAGAACCCAC 924
Db 261 MetleuThrSerGlYArYsSerhlslleuInlYserSerSerArpProHis 277

```

RESULT 2

UCK1_MOUSE	STANDARD.	PRT.	277 AA.
ID_UCK1_MOUSE			
AC_P52623			
DT_01-OCT-1996	(Rel. 34, Created)		
DT_28-FEB-2003	(Rel. 41, Last sequence update)		
DT_15-SEP-2003	(Rel. 42, Last annotation update)		

DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Krausner R.D., Collins F.S., Wagner L., Stemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusita K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raba S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mallaly S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimmett J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmechel A., Schein U.E., Jones S.J.W., Marra W.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97108719; PubMed=8951040;
RA Ropp P.A., Traut T.W.,
RT "Cloning and expression of a cDNA encoding uridine kinase from mouse
RT brain.";
RL Arch. Biochem. Biophys. 336:105-112 (1996).
CC -!- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, and N(4)-anisoylcytidine (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -!- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -!- PATHWAY: Pyrimidine salvage pathway.
CC -!- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: BC025146; AA025146.1; -.
DR EMBL: L31783; AAB50568.1; -.
DR MGD: MGI:98804; Umpk.
DR InterPro: IPR006083; PRK_UK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK_1.
DR PRINTS: PR00988; URIDINKINASE.
DR TIGRFAMs: TIGR00235; udk_1.
KM Transferase; kinase; ATP-binding.
FT NP_BIND 30 ATP (POTENTIAL).
SQ SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;

Alignment Scores:


```

Pred. No.: 6,46e-37 Length: 277
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.40% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-1 (1-1624) x UCK1_MOUSE (1-277)

QY 391 ACCTGTAAGACATCGTGGAGGCGGAAACGCTGGAGGTGCCACCTATGATTTTGACA 450
    |||||
Db 100 ThrleuYasnlleValGluglyLysThrValGlValProThrTyraeppheValThr 119
    |||||

QY 451 CACTCAAGCTTACGAGACGACGAGTGTCTACCCCTGCGAGCGTTCGTTGTTGAGCGC 510
    |||||
Db 120 HisSerArgleuProGluThrValValTyProAlaSpValValLeupheGly 139
    |||||

QY 511 ATCTGTGTTCTTAC 525
    |||||
Db 140 IleleuValPheTy 144
    |||||

RESULT 3
UCK2_HUMAN
ID UCK2_HUMAN STANDARD; PRT; 261 AA.
AC Q9BZX2; Q96KG5;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norder A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT human uridine-cytidine kinases."
RT Mol. Pharmacol. 59:1181-1186(2001).
[2]
SEQUENCE OF 15-261 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=21385121; PubMed=11494055;
RA Koizumi K., Shimamoto Y., Azuma A., Mataya Y., Matsuda A., Sasaki T.,
RA Fukushima M.;
RT "Cloning and expression of uridine/cytidine kinase cDNA from human
RT fibrosarcoma cells."
RT Int. J. Mol. Med. 8:273-278(2001).
[1]
FUNCTION: Phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate. Does not phosphorylate
deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
as a phosphate donor. Can also phosphorylate cytidine and uridine
nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
methylcytidine, and N(4)-anisoylcytidine.
[1]
CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
[1]
CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
[1]
PATHWAY: Pyrimidine salvage pathway.
[1]
TISSUE SPECIFICITY: Expressed in placenta.
[1]
SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

DR EMBL; AF236637; AAK14053.1; -
DR EMBL; AB062451; BAB56162.1; -
DR Genbank; HGNC:12562; UMPK.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK1.
DR PRINTS; PR00478; PHRIKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk.1.
KW Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29299 MW; 71791346F091EBFD CRC64;

Alignment Scores:
Pred. No.: 1.39e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 1 Gaps: 0

US-09-896-522-1 (1-1624) x UCK2_HUMAN (1-261)

QY 658 ACCTGTAAGCGCGCTTCGAGAGTCTGCTGCGACGAAGATGCGATGNG 717
    |||||
Db 187 ThrPheValValProAlaPheGluGluPheCysLeuProThrLysTyraIaSpVal 206
    |||||

QY 718 ATCATCCACGAGGA 732
    |||||
Db 207 IleleuProArgGly 211
    |||||

RESULT 4
UCK2_MOUSE
ID UCK2_MOUSE STANDARD; PRT; 261 AA.
AC Q99PM9;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last sequence update)
DE Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCK 2) (Uridine
GN monophosphokinase 2) (Cytidine monophosphokinase 2).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=21203813; PubMed=11306702;
RA Van Rompay A.R., Norder A., Linden K., Johansson M., Karlsson A.;
RT "Phosphorylation of uridine and cytidine nucleoside analogs by two
RT human uridine-cytidine kinases."
RT Mol. Pharmacol. 59:1181-1186(2001).
[1]
FUNCTION: Phosphorylates uridine and cytidine to uridine
monophosphate and cytidine monophosphate. Does not phosphorylate
deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
as a phosphate donor. Can also phosphorylate cytidine and uridine
nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
methylcytidine, and N(4)-anisoylcytidine (by similarity).
[1]
CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
[1]
CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
[1]
PATHWAY: Pyrimidine salvage pathway.
[1]
SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; AF236636; AAK14052.1; -.
DR MGD; MGI:1931744; Uck2.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; uck; 1.
KM Transferase; Kinase; ATP-binding.
FT NP BIND 27 34 ATP (POTENTIAL).
SQ SEQUENCE 261 AA; 29404 MW; 780AA3CF5CA8153 CRC64;

Alignment Scores:
Pred. No.: 1,39e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x UCK2_MOUSE (1-261)

Cy 658 ACCTTCGTAAGCGCGCTTGAGAGAGTTCGCTGCGACAAAGATAGCCGATGTG 717
Db 187 ThreVallylsProAlaPheGluGluPheCysLeuProTrpIlyTyAlaAspVal 206

Qy 718 ATCATCCACGAGGA 732
Db 207 IlleIleProArgIly 211

RESULT 5
ID UCK_DROME STANDARD; PRF; 260 AA.
ID UCK_DROME STANDARD; PRF; 260 AA.
AC Q9VC99;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE monophosphokinase) (Cytidine monophosphokinase).
GN CG6364.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil U.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

```

```

RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milentina N.V., Mobarry C., Morris J., Mosnrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuske D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster";
RL Science 287:1215-1219(2000).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003747; AAF56274.1; -.
DR FlyBase; FBgn0039179; CG6364.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
KM Hypothetical protein; Transferase; Kinase; ATP-binding.
FT NP BIND 34 41 ATP (POTENTIAL).
SQ SEQUENCE 260 AA; 29213 MW; 2E557423704925E6 CRC64;

Alignment Scores:
Pred. No.: 2.03e-06 Length: 260
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.80% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x UCK_DROME (1-260)

Qy 481 TACCGCGAGCGTGTGCTTGAGAGAGTTCGCTGCGACAAAGATAGCCGATGTG 525
Db 135 TyProlaAspValValLeuPheGluGlyIleLeuValPheTyR 149

Search completed: November 25, 2003, 07:54:01
Job time : 19.8478 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:40:49 ; Search time 76.9715 Seconds

(without alignments)
10889.155 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 536
Sequence: 1 gtcggggctgcgcctccgaccc.....ccagcttctgaggggcag 1624

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 segs, 258052604 residues

Word size: 15

Total number of hits satisfying chosen parameters: 6

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09896522/runat.21112003.184144.3228/app_query.fasta.1.2830
-DB=SPREMBL.23 -QFMT=fastan -SUFFIX=n2p_ol1.rspt -MISMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.QCGN.1.182.@runat.21112003.184144.3228 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- 1: SPREMBL_23:*
- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	169	31.5	201	4 Q96BJ0	Q96bj0 homo sapien

2	25	4.7	105	11 Q9QY8	Q9QY8 ratcus norv
3	25	4.7	111	4 Q9BU42	Q9BU42 homo sapien
4	25	4.7	261	11 Q8C476	Q8C476 mus musculu
5	15	2.8	111	4 Q92528	Q92528 homo sapien
6	15	2.8	260	5 Q8MRJ1	Q8MRJ1 drosophila

ALIGNMENTS

RESULT 1

Q96BJ0 PRELIMINARY; PRT; 201 AA.
ID Q96BJ0
AC Q96BJ0;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (0CT-2001) to the EMBL/GenBank/DDJ databases.
DR EMBL; BC015547; AAH15547.1; -.
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF0485; PRK_1.
DR PRINTS; PR00988; URIDINKINASE.
KW kinase.
SQ SEQUENCE 201 AA; 22790 MW; 0B5F2F00F7B5363 CRC64;

Alignment Scores:

Pred. No.: 5.61e-165 Length: 201
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.53% Indels: 0
DB: Gaps: 0

US-09-896-522-1 (1-1624) x Q96BJ0 (1-201)

QY	94	ATGGCTTCGGCGGAGCGCAAGACTCCGACCCCGCGAGCCGACCGTCCGCAC	153
DB	1	MetAlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluValAspArgProHis	20
QY	154	CAGCGGCTTCCTGATGAGGTGAGCGGCGGACCTGCGGAGGAGTCCGCGTGT	213
DB	21	GlnArgProPheuulIeGlyValSerGlyThrAlaSerGlyLysSerThrValCys	40
QY	214	GAGAGATCATGAGTTCCTGCGGAGCAAGACGAGTGGAACAGCGGACCGGAGGTGTC	273
DB	41	GluValIleMetGluLeuLeuGluGlnAsnGluValGluGlnArgGlnArgValVal	60
QY	274	ATCTGAGCCAGGACGAGTTCACAGGCTTCGAGCGGACGAGCAAGAGGCTTGG	333
DB	61	IleLeuSerGlnAspArgPheThrLysValLeuThrAlaGluGlnLysAlaLeu	80
QY	334	AAAGGACATGACATTTTGGACATCCAGATGCTTGAATGAATTTGATGACAGGACT	393
DB	81	LysGlyGlnThrAsnPheAspHisProAspAlaPheAspAsnLeuMetHisArgThr	100
QY	394	CTGAGAGACATGCTGGAGGCGCAAAACGCTGAGTGTCGACCATGATTTTGGACACAC	453
DB	101	LeuLysAsnIleValGluGlnLysThrValGluValProThrTyrAspPheValThrHis	120
QY	454	TCAAGTTTCCAGAGACACCGGTGCTTACCTCTGCGGACGTCGTCTGTTGAGGCATC	513
DB	121	SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle	140
QY	514	TTGGTGTCTTACAGCAGAGATCCGGGACATGTTCCACCTGCGGCTCTTGAGCACAC	573

```
DB 141 LeuValPheTyrSerGlnGluIleArgAspMetCehiSleuArgLeuPheValAspThr 160
QY 574 GACTCCGACGCTGAGCTGTCTCGAGA 600
DB 161 AsperAepValArgLeuSerArgArg 169

RESULT 2
Q9QY68 PRELIMINARY; PRT; 105 AA.
ID Q9QY68
AC Q9QY68
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Uridine Kinase (Fragment).
GN UK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD; TISSUE=Spinal cord;
RX MEDLINE=20050059; PubMed=10581173;
RA Yuh I., Yaoi T., Matsunabe S., Okajima S., Hirasawa Y., Fushiki S.;
RT "Up-regulated uridine kinase gene identified by RUCS in the ventral
RT horn after crush injury to rat sciatic nerves.";
RL Biochem. Biophys. Res. Commun. 266:104-109(1999).
DR EMBL; AB030700; BAA83085.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR KINASE.
FT NON_TER 1 105
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11967 MW; 9DB43C15E07EB029 CRC64;

Alignment Scores:
Pred. No.: 3.66e-16 Length: 105
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-1 (1-1624) x Q9QY68 (1-105)
QY 658 ACCTTGTAAGCGGCGCTTCGAGAGTTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 36 ThirPheValIysProAlaPheGlnGluPheCysLeuProThrIlystYrAlaAspVal 55
QY 718 ATCATCCCAAGAGA 732
DB 56 IleIleProArgIly 60

RESULT 3
Q9BU42 PRELIMINARY; PRT; 111 AA.
ID Q9BU42
AC Q9BU42
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to uridine monophosphate kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strauberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC002906; AAH02906.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR KINASE.
SQ SEQUENCE 111 AA; 12587 MW; E6688B1B86F432A9 CRC64;

Alignment Scores:
Pred. No.: 3.64e-16 Length: 111
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 4 Gaps: 0

US-09-896-522-1 (1-1624) x Q9BU42 (1-111)
QY 658 ACCTTGTAAGCGGCGCTTCGAGAGTTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 37 ThirPheValIysProAlaPheGlnGluPheCysLeuProThrIlystYrAlaAspVal 56
QY 718 ATCATCCCAAGAGA 732
DB 57 IleIleProArgIly 61

RESULT 4
Q8C476 PRELIMINARY; PRT; 261 AA.
ID Q8C476
AC Q8C476
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Uridine-cytidine Kinase 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The Riken Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082837; BAC38646.1; -.
SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:
Pred. No.: 3.27e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 11 Gaps: 0

US-09-896-522-1 (1-1624) x Q8C476 (1-261)
QY 658 ACCTTGTAAGCGGCGCTTCGAGAGTTCTGCTGCCGACAAGAAGTATGCCGATGTG 717
DB 187 ThirPheValIysProAlaPheGlnGluPheCysLeuProThrIlystYrAlaAspVal 206
QY 718 ATCATCCCAAGAGA 732
DB 207 IleIleProArgIly 211

RESULT 5
Q92528 PRELIMINARY; PRT; 111 AA.
ID Q92528
AC Q92528
DT 01-FEB-1997 (TREMBlrel. 02, Created)
```

DT 01-FEB-1997 (TrEMBLrel. 02, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE 5'-terminal region of UMK.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=96411689; PubMed=8812458;
 RA Ozaki K., Kuroki T., Hayashi S., Nakamura Y.,
 RT "Isolation of three testis-specific genes (TSA303, TSA806, TSA903) by
 RT a differential mRNA display method.";
 RL Genomics 36:316-319 (1996).
 DR EMBL: D78335; BAA1349.1;
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 SQ SEQUENCE 111 AA; 12617 MW; 81F6E4019CF428AB CRC64;

Alignment Scores:
 Pred. No.: 7.81e-06 Length: 111
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.80% Indels: 0
 DB: 4 Gaps: 0

US-09-896-522-1 (1-1624) x Q92528 (1-111)

QY 658 ACCTTGCGAAGCGGCTTCGAGAGCTTCGCTGCCGCAAG 702
 DB 37 ThPheValtyProAlaPheGluGluPheCysLeuProThrIys 51

RESULT 6

08MRJ1 PRELIMINARY; PRT; 260 AA.
 AC 08MRJ1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE LD13909p.
 GN CG6364.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Stapleton M., Broksstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Fise B.,
 RA George R., Gonzalez W., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclab U., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.B., Rubin G.M.,
 RA Celisner S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AY119583; AA550237.1; -;
 DR FlyBase: FBgn0039179; CG6364.
 DR InterPro: IPR006083; PRK_URK.
 DR InterPro: IPR000764; Uridine_kin.
 DR Pfam: PF00485; PRK; 1.
 DR PRINTS: PR00988; URIDINKINASE.
 SQ SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A1655 CRC64;

Alignment Scores:

Pred. No.: 7.03e-06 Length: 260
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.80% Indels: 0
 DB: 5 Gaps: 0

US-09-896-522-1 (1-1624) x Q8MRJ1 (1-260)

QY 481 TACCCTGCGAGCGTGTCTGTTTGAGGCGCATCTTGNGTCTAC 525
 DB 135 TyrProAlaAspValValLeuPheGluGluIleLeuValPheTyr 149

Search completed: November 25, 2003, 07:58:08
 Job time: 79.9715 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 07:43:14 ; Search time 30.7225 Seconds
(without alignments)
10166.994 Million cell updates/sec

Title: US-09-896-522-1
Perfect score: 536
Sequence: 1 gtggggtgcgcgcgcaccc.....ccaggtgtgtggggccag 1624

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0896522/runat_21112003_184145_3234/app_query.fasta_1.2830
-DB=PIR_76 -QFMT=fastan -SUFFIX=n2p_ol1.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR SCORE=quality -THR MIN=15 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0896522.@CGN_1_1_68.@runat_21112003_184145_3234 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
No matches found					

Search completed: November 25, 2003, 07:59:52
Job time : 30.7225 secs

This Page Blank (uspto)

..... (uspto)

PT such as central nervous system

Description
Human poly(phenylacetylene)
Human polypropylene
Human protein sequencing
Human, 57658 protease
Protein of a humane
Human uridine kinase
Human polypropylene
Human MDR1 protein
Protein relating to
Human poly(phenylacetylene)
Human transferrase
Human prostatic cancer
Human ovarian anti
Human musculoskeletal
Novel human musculin
Human 57658 protease
Drosophila melanogaster
Human poly(phenylacetylene)

XX Example 2; SEQ ID NO 6219; 10078bp; English.

PS The invention relates to human nucleic acids (AA157798-AA161369) and

XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytosstatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, Leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

SO Sequence 296 AA;

Alignment Scores:

Pred. No.:	1,3e-295	Length:	296
Score:	296.00	Matches:	296
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.22%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-1 (1-1624) x AAM41288 (1-296)

QY 37 GCCCGGGGGGGGAAAGGGGGGGGGGAGCCCGATGCCGGGAGCGGAGCCCGAGATG 96

Db 1 AAlaGlyProGlyGluGlyAlaGlyAlaGlyThrArgCysAlaGlyAlaGluAlaGluMet 20

QY 97 GCTTCGGCGGGAGCGGAAAGACTGCGAGAGCCCGCGCGCGAGCGGAGCCGTCGACAG 156

Db 21 AlaSerAlaGlyGlyGluAspCysGluSerProAlaProGluAlaAspArgProHisGln 40

QY 157 CGGCGCTTCTGATAGGGGTGAGCGCGCGCACTGCCAGCGGAAAGTCGACCGTGTGAG 216

Db 41 ArgProHeuLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCysGlu 60

QY 217 AAGATCATGAGTGTCTGCGGAGCAGAAACGAGGTGGAACGCGGAGCCGGAAGTGTATC 276

Db 61 LysIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValIle 80

QY 277 CTGAGCCAGGACAGGTTCTACAGAGTCTGACGGGAGAGCAGAGCGGACCTTGAAA 336

Db 81 LeuSerGlnAspArgPheTrpLysValLeuThrAlaGluGlnLysValLysAlaLeuLys 100

QY 337 GGACAGTACATTTTGAACATCCAGATGCCCTTGTATATGATTTGATGACAGGACTCTG 396

Db 101 GlyIleTrpAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThrLeu 120

QY 397 AAGAACTCCGCGGAGCGGAGAAACGCTGAGTGGCGGACCTATGATTTTGTGACACTCA 456

Db 121 LysAsnIleValGluGlyLysThrValGluValProThrTrpAspPheValThrHisSer 140

QY 457 AGGTACAGAGACACGAGTGTCTACAGCTGCGGAGCTGCTGTGTTGAGGAGCATCTTG 516

Db 141 ArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIleLeu 160

QY 517 GTGTTTACAGCCGAGAGATCCGGACATGTTCCAGCTGCGCTCTTCTGTGACACCGAC 576

Db 161 ValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThrAsp 180

QY 577 TCCGACGTGAGGCTGTCTCGAAGAGTTCTCCGGAGCGTGGCCGAGGAGGAGACTTGAG 636

Db 181 SerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeuGlu 200

QY 637 CAGATTCTGACGACGTAACACACCTTGCTGTAAGCGCGGCTTCGAGGAGTTCTGCTCGC 696

Db 201 GlnIleLeuThrGlnTrpThrPheValLysProAlaPheGluLysIlePheCysLeuPro 220

QY 697 ACAAGAAGTATGCCATGTATCATCCACGAGAGTGGACAAATGTGTCATCAAC 756

Db 221 ThrLysLysTyrAlaAspValIleIleProArgGlyValAlaAspMetValAlaIleAsn 240

QY 757 CTGATGTCGAGACATCTCCAGGACATTTGAAATGGAGCATCTGCAATGGGACCGAGGA 816

Db 241 LeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysTrpHisArgGly 260

QY 817 GGATCCAAATGGCGGAGCTTACAAGCGGACCTTTTCTGAGCCAGGAGGAGCCCTGGAGAT 876

Db 261 GlySerAsnGlyArgSerTyrLysArgThrPheSerGluProGlyAspHisProGlyMet 280

QY 877 CTGACCTTCTGGCAACGCTCACATTTGGAGTCCAGAGCAGACCCAC 924

Db 281 LeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 296

RESULT 2

AAM39502

ID AAM39502 standard; Protein; 277 AA.

XX

AC AAM39502;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 2647.

XX

XX Human; nootropic; immunosuppressant; cytosstatic; gene therapy; cancer;

KW peripheral nervous system neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW Leukaemia.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR N-PDB; AA158658.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

XX such as central nervous system injuries -

XX

Example 4; SEQ ID NO 2647; 10078bp; English.

XX

CC The invention relates to human nucleic acids (AA157798-AA161369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytosstatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

DR WPI; 2002-140091/18.
 DR N-PSDB; AAD27186.
 XX
 PT New isolated human uridine kinase family polypeptide 57658, useful for
 PT treating hematopoietic neoplastic disorders and disorders of neurons,
 PT heart and blood vessels
 XX
 PS Claim 4; Fig 1a; 103pp; English.
 XX
 CC The patent discloses human uridine kinase-like polypeptides, designated
 CC 57658 and polynucleotides encoding such proteins. 57658 DNAs and proteins
 CC are useful for developing diagnostic and therapeutic agents for 57658-
 CC mediated or related disorders such as haematopoietic neoplastic disorders
 CC (e.g. leukaemia), haematoid disorders (e.g. psoriasis, dermatitis,
 CC diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets,
 CC osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels
 CC (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's
 CC disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or
 CC their antibodies are useful in screening assays, detection assays (e.g.
 CC forensic biology) and predictive medicine (e.g. diagnostic assays).
 CC prognostic assays, and monitoring clinical trials and pharmacogenomics).
 CC They are useful as reagents for diagnosing and treating 57658-mediated
 CC disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping,
 CC to identify an individual from a minute biological sample (tissue typing)
 CC and to aid in forensic identification of the biological sample. The
 CC present sequence is human 57658 protein.

XX Sequence 277 AA;

Alignment Scores:

Pred. No.:	4,91e-276	Length:	277
Score:	277.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	51.68%	Indels:	0
DB:	23	Gaps:	0

US-09-896-522-1 (1-1624) x AAB16592 (1-277)

QY 94 ATGGCTTCGGGGAGGCGGAAGCTCGAGAGCCCGCGCGGACCGGACCTCCGAC 153
 DB 1 MetAlaSerIaGlyGluAerCysGluSerProIaProGluIaAaAraGProHis 20
 QY 154 CACGGCCCTTCCTGATAGGGGTGAGCGGGGCACTCCGCGGAAGTGAACCTGTGT 213
 DB 21 GlnArGProHeuIleGlyValSerGlyGlyThrIaSerGlyLysSerThrValCys 40
 QY 214 GAGAAATCATGAGTTGCTGAGCAGAAAGAGTGGAGCAGCGGCGGAGAGTGTGTC 273
 DB 41 GluLysIleMetCtIuLeuLeuGluGlnAerGluValGluGlnIaGlyValVal 60
 QY 274 ATCTTAGAGCAGAGCAGGTTCTACAAAGTCTTGAACGGCAGAGCAAGAGCCCTTG 333
 DB 61 IleLeuSerCtIaAspArGpHeuTyrLysValIleuThrIaGluGlnValIaLysAlaLeu 80
 QY 334 AAGGACAGTAACTTTTGCATCCAGATGCTTTGATATGATTGATGACAGACT 333
 DB 81 LysGlyGlnTyrAsnPhaAspHisProAspAlaPhaAspAsnAspLeuMetHisArgThr 100
 QY 394 CTGAAGAATCATGCTGAGGCGCAAAACGATGAGGTGCGGACTATGATTGTCACAC 453
 DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrIaAspPheValThrHis 120
 QY 454 TCAAGTTACAGAGACACGAGTGTTCACCTCGCGGACGTGTCTGTTGAGGCGATC 513
 DB 121 SerArgLeuProGluThrIaThrValIaLysProAlaAspValIaLeuPheGluGlyIle 140
 QY 514 TTGGTGTTCACAGCGCAGAGATCCGGGACATGTCACCTGGCCCTTCGTGGACACC 573
 DB 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTCTCCGGGACCTGGCGCAGGAGGAGACTG 633

DB 161 AspSerAspValArgLeuSerArgValIleuArgAspValArgArgGlyArgAspLeu 180
 QY 634 GAGCAGATTCTGACGAGTGAACCACTTCGTGGAAGCGGCTTCGAGAGTTCGCTG 693
 DB 181 GluGlnIleuThrIaThrIaThrIaPheValLysProAlaPheGluGluPheCysLeu 200
 QY 694 CCGCAAGAAGATGCCGATGATCATCCACGAGAGTGGACAAATGTTGTCATC 753
 DB 201 ProThrLysLysTyrIaAspValIleIleProArgGlyValaAspAsnMetValaIaIle 220
 QY 754 AACCTGATCGTGCAGACATCCAGACATTTGAAATGTGACATCTGCAATGCAACGA 813
 DB 221 AsnIleValGlnIleHisIleGlnAerIleLeuAsnGlyAspIleCysLysTrpHisArg 240
 QY 814 GGAGGTCCAAATGGGCGGAGCTTCAAGCGGACCTTTTTCGAGCCAGGAGCAACCTGG 873
 DB 241 GlyGlySerAsnGlyArgSerTyrLysArgGlnPheSerGluProGluAspHisProGly 260
 QY 874 ATGCTGACCTTCGCAAAAGCTTCACATTTGAGTCCAGACAGACCCAC 924
 DB 261 MetLeuThrSerCtLysArgSerHisLeuGluSerSerArgProHis 277
 RESULT 5
 ID AA014412
 XX AA014412 standard; Protein; 260 AA.
 AC AA014412;
 XX
 DT 02-MAY-2002 (first entry)
 XX
 DE Protein of a human uridine kinase (UDK).
 XX
 KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
 KW probe; chromosome localisation study; tissue expression; gene therapy;
 KW antibody; vaccine; human ovarian cancer; immunological disorder;
 KW human colon carcinoma; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200172963-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09663.
 XX
 PR 27-MAR-2000; 2000US-0536647.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Ho YS, Johnson RK;
 XX
 DR WPI; 2001-626259/72.
 DR N-PSDB; AAK98735.
 XX
 PT Novel human uridine kinase polypeptides useful for treating cancers,
 PT and to identify agonists and antagonists of the polypeptide useful for
 PT treating conditions associated with uridine kinase imbalance
 XX
 PS Claim 3; Page 29-30; 31pp; English.
 XX
 CC The invention relates to newly identified human uridine kinase (UDK)
 CC polypeptides and polynucleotides and methods for producing such
 CC polypeptides by recombinant techniques. Also disclosed in the invention
 CC are methods for utilising uridine kinase polypeptides and polynucleotides
 CC in diagnostic assays. The polynucleotides and polypeptides of the
 CC invention may be used as diagnostic reagents by detecting mutations in an
 CC associated gene. An array of oligonucleotide probes comprising the
 CC uridine kinase polynucleotide sequence or fragments thereof can be
 CC constructed to conduct efficient screening of genetic mutations, for
 CC example. Detection of abnormally decreased or increased levels of
 CC polypeptide or mRNA expression may also be used for diagnosing or
 CC determining susceptibility of a subject to a disease of the invention.
 CC The polynucleotide sequences of the invention can be used for chromosome

CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinomas, and
CC immunological disorders. This sequence represents the protein of a human
CC uridine kinase of the invention.

XX SQ Sequence 260 AA;

Alignment Scores:
Pred. No.: 1.62e-258 Length: 260
Score: 260.00 Matches: 260
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.51% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-1 (1-1624) x AAO14412 (1-260)

QY 94 ATGGCTTCGGGAGGAGGAGAGAGTGGAGAGCCCGGCGGAGCGGACCGTCCGAC 153
DB 1 MetalaSerAlaGlyGlyAlaSPcylUserProAlaProGluAlaAspArgProHis 20
QY 154 CAGCGGCGCTTCGATAGGGGTGAGCGGCGGCACTGCGAGCGGAAAGTCAGCCGTGT 213
DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY 214 GAGAGATCATGAGTCTCTGGACAGACAGAGGTGAAACGCGGACGCGGAGAGTGT 273
DB 41 GluLysIleMetGluLeuGluGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
QY 274 ATCTGAGCCGAGGACAGGTTTACAGAGTCTGACGAGAGGAGGAGGCGCAAGCGCTTG 333
DB 61 IleuSerGlnAspArgPheTyrlValValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTACAAATTTTGACATCCAGATCCATTGATATGATTTGATGACAGAGCT 393
DB 81 LysGlyGlnTyrlAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAAGAAATCGTGGAGGGGAAACGGTGGAGGTCCGACCTTGTATTTGTGACACAC 453
DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrlAspPheValThrHis 120
QY 454 TCAAGTTACAGAGACACGCGTGTCTACCCCTGCGAGCGGTCTGTTGAGGGGATC 513
DB 121 SerArgLeuProGluThrValValValTyrlProAlaAspValValLeuPheGluGlyLe 140
QY 514 TTGGTGTCTTACAGCAGAGATCCGGGACATGTTCCACCTGCGCTTCTGTGACACC 573
DB 141 LeuValPheTyrlSerGlnGluIleArgAspPhePheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGACGTCGCGCGAGGGAGGACCTG 633
DB 161 AspSerAspValArgLeuSerArgValValLeuArgAspValArgArgGlyArgAspLeu 180
QY 634 GAGGAGTTCTGAGCAGTACACACCTGCTGAAGCGCGCTTCGAGGAGTTCTGCTG 693
DB 181 GluGlnIleLeuThrGlnTyrlThrThrPheValLysProAlaPheGluGluPheCysLeu 200
QY 694 CCGACAAAGAGTATGCGATGTGATCATCCACGAGAGTGAACAATATGTTGCTCATC 753
DB 201 ProThrLysLysTyrlAlaAspValIleIleProArgGlyValAlaAspMetValAlaIle 220
QY 754 AACCTGATCGTGACGACATTCAGGACATTTGATGTTGATCATCTGCAATATGGACCGA 813
DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrHisArg 240
QY 814 GAGAGGTCCATGGGCGGAGGTAACAAGCGGCTTTCTGAGCCAGGGGACACCCCTGGG 873
DB 241 GlyGlySerAsnGlyArgSerTyrlLysArgThrPheSerGluProGlyAspHisProGly 260

RESULT 6
ID AAG64506
AC AAG64506; standard; Protein; 276 AA.

DT 02-OCT-2001 (first entry)

XX Human uridine kinase.

XX Human; uridine kinase; UK.

XX Homo sapiens.

XX CN1287172-A.

XX 14-MAR-2001.

XX 07-SEP-1999; 99CN-0118818.

XX 07-SEP-1999; 99CN-0118818.

XX (YTFU-) UNIT FUDAN.

XX Yu L, Zhao Y, Zhang H;

XX WPI: 2001-409529/44.

XX N-PSDB; AAH75355.

XX Human uridine kinase and its coding sequence, preparation and

XX application -

XX Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.

XX The invention relates to human uridine kinase (UK).

XX SQ Sequence 276 AA;

Alignment Scores:
Pred. No.: 2.28e-185 Length: 276
Score: 189.00 Matches: 189
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.26% Indels: 0
DB: 22 Gaps: 0

US-09-896-522-1 (1-1624) x AAG64506 (1-276)

QY 265 AAGGTGTCATCTGAGCCAGACAGGTTCTACAGATCCTGACGACAGAAAGCC 324
DB 57 LysValValIleLeuSerGlnAspArgPheTyrlValValLeuThrAlaGluGlnLysAla 76
QY 325 AAGGCTTGAAGAGACATCAATTTTGACATCCAGATGCTTGTATATGATTTGATG 384
DB 77 LysAlaLeuLysGlyGlnTyrlAsnPheAspHisProAspAlaPheAspAsnAspLeuMet 96
QY 385 CACAGAGCTCTGAAGAAATCGTGGAGGCAAAACGGTGGAGGTGCGACTATGATTTT 444
DB 97 HisArgThrLeuLysAsnIleValGluGlyLysThrValGluValProThrTyrlAspPhe 116
QY 445 GTGACACTCAAGTTTACAGACAGACACGATGCTTACCTGCGGACCTGGTTCTGTT 504
DB 117 ValThrHisSerArgLeuProGluThrThrValValTyrlProAlaAspValValLeuPhe 136
QY 505 GAGGGCATCTTGGTGTCTTACAGCCAGAGATTCGGGACATGTTCCACTGGCGCTTTC 564
DB 137 GluGlyIleLeuValPheTyrlSerGlnGluIleArgAspPhePheHisLeuArgLeuPhe 156
QY 565 GTGACACGACGTCGAGCGTCAAGGCTGTCTCGAAGAGTTCTCCGGGACGTCGCGAGGG 624
DB 157 ValAspThrAspSerAspValArgLeuSerArgValValLeuArgAspValArgArgGly 176
QY 625 AAGGACCTGAGAGATTTCTGACGACATCACACCTTCTGTGAAGCGGCTTTCGAGAG 684

Db 177 ArgAspLeuInGlnLeuThrInGlnTyr-ThrThrPheValLysPheAlaPheGln 196
Qy 685 TTCTGCTGCGCCGAAAGATATGCCGATGTATCATCCACAGAGATGACATATG 744
Db 197 PheCysLeuProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMet 216
Qy 745 GTTGCCATCAACCGATCGTGACACATCATCTGGAATGGTGAATCTGCAAA 804
Db 217 ValAlaIleAsnLeuIleValGlnHisIleGlnHisPheLeuAsnGlyAspIleCysLys 236
Qy 805 TGGCACCAGAGAGGTCCTCAATGGCGG 831
Db 237 TrpHisArgGlyGlySerAsnGlyArg 245
RESULT 7
ABBB9353
ID ABB89353 standard; Protein; 190 AA.
AC ABB89353;
XX
XX 24-MAY-2002 (first entry)
DE Human polypeptide SEQ ID NO 1729.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW valineary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX 29-NOV-2001.
PD
XX
PF 18-MAY-2001; 2001MO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
DR N-PSDB; ABL89762.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
XX Claim 11; SEQ ID NO 1729; 2081tp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 190 AA;
Alignment Scores:
Pred. No.: 9,58e-131 Length: 190
Score: 136.00 Matches: 136
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25,37% Indels: 0
DB: 23 Gaps: 0
US-09-896-522-1 (1-1624) x ABB89353 (1-190)
Qy 193 AGCGGGAGTCGACCGCTGTGTGAGAGATCATGAGTTGCTGGACAGAACGAGTGGAA 252
Db 34 SerGlyLysSerThrValCysGlnLysIleMetGlnLeuLeuValGlnValGln 53
Qy 253 CAGCGGACGCGGAGGAGGTGTATCTCTGAGCCAGACAGGTTCTACAAAGTCTGACGGCA 312
Db 54 GlnArgGlnArgGlyValValIleLeuSerGlnAspArgPheTyrIleValLeuThrAla 73
Qy 313 GAGCAGAGGCGCAAGGCTTGAAGGACAGTACATTTTGACCATCCAGATGCTTTGAT 372
Db 74 GlnGlnLysAlaIleValLeuLysGlnTyrAsnPheAspHisProAspAlaPheAsp 93
Qy 373 AATGATTTGATGACAGCACTCTGAGAAATCTGAGGCAAAACGGTGGAGTGGCGG 432
Db 94 AsnAspLeuMetHisArgThrLeuLysAsnIleValGlnGlyLysThrValGlnValPro 113
Qy 433 ACCTAATTTTGTGACACACTCAAGTTACAGACAGACAGCGGTGTCTTACCTGGCGGAC 492
Db 114 ThrTyrAspPheValThrIleSerArgLeuProGlnThrThrValValTyrProAlaAsp 133
Qy 493 GTGTTCTGTGTTGAGGGCATCTTGTCTTACAGCCAGAGAAATCCGGACATGTTCCAC 552
Db 134 ValValLeuPheGlnGlyIleLeuValIlePheTyrSerGlnGlnIleArgAspMetPheHis 153
Qy 553 CTGCGGCTCTCTGTGTGACACCGACTCCGACGTCGACGCTGTCCGAGA 600
Db 154 LeuArgLeuPheValAspThrAspSerAspValArgLeuSerArgArg 169
RESULT 8
ABG70328
ID ABG70328 standard; Protein; 120 AA.
XX
XX AC ABG70328;
XX
XX 21-OCT-2002 (first entry)
DE Human MDDT protein Incyte ID No: LI:235557.12.orf2:2001JAN12.
XX
XX KW Human; molecule for disease detection and treatment; MDDT; cancer;
XX cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
XX autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antisthmatic.
XX
XX OS Homo sapiens.
XX
XX PN WO200255738-A2.
XX
XX 18-JUL-2002.
XX
XX PD 09-JAN-2002; 2002MO-US01008.
XX
XX PF 12-JAN-2001; 2001US-261622P.
XX
XX PR 16-JAN-2001; 2001US-261865P.
XX
XX PR 17-JAN-2001; 2001US-262208P.
XX
XX PR 17-JAN-2001; 2001US-262209P.
XX
XX PR 17-JAN-2001; 2001US-262326P.
XX
XX PR 19-JAN-2001; 2001US-263063P.
XX
XX PR 19-JAN-2001; 2001US-263065P.
XX
XX PR 19-JAN-2001; 2001US-263329P.

XX	(INCY - INCYTE GENOMICS INC.
PA	
XX	Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AX,
PI	Dam TC, Liu TP, Harris B, Flores V, Dafio A, Marwaha R, Chen AJ;
PI	Chang SC, Gerstein EH, Peralta CH, David MH, Lewis SA;
XX	WPI; 2002-590679/63.
DR	N-Psdb; AB51801.
PT	New disease detection and treatment molecule (MDR) polynucleotides and
PT	polypeptides, useful in diagnosing, studying, preventing or treating
PT	diseases associated with MDR expression, e.g. autoimmune or
XX	inflammatory disorders -
PS	Claim 27; Page 120; 129pp; English.
XX	
CC	The present invention relates to the isolation of novel human
CC	molecules for disease detection and treatment (MDR), and the
CC	polynucleotide sequences (mdr) encoding them. The MDR polypeptides
CC	may be used to screen for molecules that bind to, or are bound by the
CC	encoded polypeptides, and to develop a transcript image of a tissue or
CC	cell type. Probes comprising at least 20 nucleotides of the mdr
CC	polynucleotide may be used to assess the toxicity of a test compound.
CC	The MDR polypeptides and mdr polynucleotides are useful in the
CC	diagnosis, study, prevention and treatment of diseases associated with
CC	the expression of cell proliferative for disease detection and treatment. Such
CC	disorders include cell proliferative disorders (e.g. arteriosclerosis,
CC	cirrhosis, or cancers), and autoimmune/inflammatory disorders
CC	(e.g. asthma, Crohn's disease, or multiple sclerosis). The mdr
CC	polynucleotides may also be used as molecule markers, in microarrays,
CC	and in somatic or germline gene therapy. ABG70306-ABG70341 represent
CC	the MDR proteins of the invention.
XX	
SQ	Sequence 120 AA;
Alignment Scores:	
Pred. No.:	5.37e-106 Length: 120
Score:	112.00 Matches: 112
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	20.90% Indels: 0
DB:	Gaps: 0
US-09-896-522-1 (1-1624) x ABG70328 (1-120)	
OY	361 GATGCCCTTTGATATGATTGGATGCACAGACTTGAAAGAACATCGTAGAGCAAAAACG 420
Db	5 Asplaphheasphendaspneuethisatgtnrlleuylsaenillevalglugllylvsthr 24
OY	421 GTGAGGTGCCGACCTTATGTTTTTGTGAACAACACTCAAGTTTCAGAGAACAACGGTGTGC 480
Db	25 ValglivalprothrlyrasphevalttrhtsserfaglsuepfrogltutThrValVal 44
OY	481 TACCCTCGGACCTGGCTTCTGTTTGGAGCATCTTGGTGTTTATACAGCAGCAATCCGG 540
Db	45 TyrProhlaaspyalValleuphegiungllyllevalPheTySergInuilearg 64
OY	541 GACATGTTCCACCTCGCGCTCTTGTGACACACCACTCCGACGTGAGGTCTCTCAGAA 600
Db	65 AspMetPhehisleuairgleuphevalAsptHraSpserAepValArgLeuSerArgArg 84
OY	601 GTTCTCCGGAGCGTGGCGGAGGAGGACCTGGAGCAGATTCTGACCGCACTACACACC 660
Db	85 ValleuargrspalVlrIargaglYIarGaBpleuGlunllleUvThrgInTYrThrThr 104
OY	661 TTGCTGAAGCCGGCTTCGAGAGAGTTCTGCGCG 696
Db	105 PheVallysrProhlaaphegiunglunPhecysleupro 116
RESULT 9	
AAOI4413	
ID	AAOI4413 standard; Protein; 277 AA.

XX AAO14413;
AC
XX
DT 02-MAY-2002 (first entry)
XX
XX
DE Protein relating to a human uridine kinase (UDK) of the invention.
XX
XX Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KM probe; chromosome localisation study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen.
XX
XX Unidentified.
OS
XX
XX WO200172963-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
XX
XX
XX Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
PT
PS Disclosure; Page 23; 31pp; English.
XX
XX
XX The invention relates to newly identified human uridine kinase (UDK)
CC polypeptides and polynucleotides and methods for producing such
CC polypeptides by recombinant techniques. Also disclosed in the invention
CC are methods for utilising uridine kinase polypeptides and polynucleotides
CC in diagnostic assays. The polynucleotides and polypeptides of the
CC invention may be used as diagnostic reagents by detecting mutations in an
CC associated gene. An array of oligonucleotide probes comprising the
CC uridine kinase polynucleotide sequence or fragments thereof can be
CC constructed to conduct efficient screening of genetic mutations, for
CC example. Detection of abnormally decreased or increased levels of
CC polypeptide or mRNA expression may also be used for diagnosing or
CC determining susceptibility of a subject to a disease of the invention.
CC The polynucleotide sequences of the invention can be used for chromosome
CC localisation studies and tissue expression studies. The polypeptides of
CC the invention or fragments thereof may be used as immunogens to produce
CC antibodies. These antibodies may be employed to isolate or identify
CC clones expressing the polypeptide. The polypeptides and polynucleotides
CC of the invention can be used as a vaccine or in gene therapy to treat
CC diseases such as human ovarian cancer, human colon carcinoma, and
CC immunological disorders. This sequence represents the protein relating to
CC a human uridine kinase (UDK) of the invention.
CC NOTE: The present sequence is stated as being the same as that shown as
CC SEQ ID NO: 2 in the sequence listing of the specification. However, the
CC sequences differ.
XX
XX
SQ Sequence 277 AA;

Alignment Scores:

Pred. NO.:	5.26e-37	Length:	277
Score:	45.00	Matches:	45
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.40%	Gaps:	0
DI:	22		

US-09-896-522-1 (1-1624) x AAO14413 (1-277)

391 ACTCTGAAGAACATCTGCTGAGGGCAGAAACGGTGCAGGTGCCGACCTATGATTTGTGACA 450

Db 100 ThrlenuysasnileValgluglyLyeThrValgluValProThrTyraSpheValThr 119
XX
QY 451 CACTCAAGGTTTACCAGACGACGAGTGTCTACCTCGGACGGTTCGTTGAGGCG 510
Db 120 HisserTrglueupProgluThrValValTyrrProAlaSpValValLeupheglugly 139
QY 511 ATCTTGSTGTTCTTAC 525
Db 140 lleuValphetyr 144
RESULT 10
AAM38694
ID AAM38694 standard; Protein; 261 AA.
XX
AC AAM38694;
XX
XX 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1839.
XX
KM Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AA157850.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
PS Example 3; SEQ ID NO 1839; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC
CC Note: The sequence data for this patent did not form part of the printed

CC specification.
XX
SQ Sequence 261 AA;
XX
Alignment Scores:
Pred. No.: 2,14e-16 Length: 261
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.66% Indels: 0
DB: 22 Gaps: 0
US-09-896-522-1 (1-1624) x AAM38694 (1-261)
QY 658 ACCTTGCTGAAGCGGCTTCGAGGAGTTCGCTCGCCGACAAAGAATATGCCGATG 717
Db 187 ThrPheVallyseProAlaPhegluglyupheCysleuProThrlystyrrAlaSpVal 206
QY 718 ATCATCCCAAGAGGA 732
Db 207 lleileProArggly 211
RESULT 11
AAB73494
ID AAB73494 standard; Protein; 261 AA.
XX
AC AAB73494;
XX
XX 31-JUL-2001 (first entry)
XX
DE Human transferase HTFS-1, SEQ ID NO:1.
XX
XX Human transferase; HTFS; agonist; antagonist; cellular signalling;
KM proliferation; cell proliferative disorder; immune disorder;
KM atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
KM cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
KM haemodialysis; extracorporeal circulation; trauma; transgenic animal;
KM gene therapy; drug screening.
XX
OS Homo sapiens.
XX
XX WO200132888-A2.
XX
XX 10-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-US30485.
XX
XX 04-NOV-1999; 99US-0163595.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
PI Shih LL, Razimzai Y, Lu DAM, Baughn MR;
XX
DR WPI; 2001-328796/34.
DR N-PSDB; AAH23801.
XX
XX Human transferase polypeptides and polynucleotides useful for
PT diagnosis, prevention and treatment of cell proliferative and immune
PT system disorders and for identifying agonists and antagonists -
XX
XX Claim 1; Page 103-104; 157pp; English.
XX
CC Sequences AAB73494-AAB7335 represent novel human transferase proteins
CC HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
CC encoding them. The proteins play important roles in the regulation of
CC cellular signalling and proliferation. The HTFS proteins are useful for
CC screening compounds for their effectiveness as agonists or antagonists of
CC transferase activity, or for compounds that specifically bind to an HTFS
CC protein or which modulates the activity of an HTFS protein.

CC Pharmaceutical compositions comprising an HTFS protein, HTFS
CC agonist or antagonist, or genetic construct encoding an HTFS
CC protein are useful for treating a disease or condition associated
CC with decreased or increased expression of functional HTFS. Disorders
CC which may be treated using such compositions include cell proliferative
CC disorders and immune disorders. For example, diseases which may be
CC treated include atherosclerosis, hepatitis, psoriasis, cancers (including
CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
CC complications of cancer, haemodialysis, extracorporeal circulation,
CC trauma and haematopoietic cancer, including lymphoma, leukaemia and
CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
CC transgenic animals to model human diseases, for diagnostic purposes and
CC to generate hybridisation probes useful in mapping the naturally
CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HTFS may be
CC used for the diagnosis of disorders associated with the expression of
CC HTFS, or in assays to monitor patients being treated with HTFS or
CC agonists, antagonists or inhibitors of HTFS. The present sequence
CC represents an HTFS protein of the invention.

XX Sequence 261 AA;

Alignment Scores:

Pred. No.:	2.14e-16	Length:	261
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.66%	Indels:	0
DB:	22	Gaps:	0

US-09-896-522-1 (1-1624) x AAB73494 (1-261)

QY 658 ACCTTGCGAAGCCGGCCTTCGAGAGTTCTGCTGCCGACAAAGATATGCCGATGTG 717
Db 187 ThrpheVallyspProAlaPheGluGluPheCysLeuProThrlyblySTyAlaAspVal 206
|||||

QY 718 ATCATCCACGACGAGA 732
Db 207 IleIleProArgGly 211
|||||

RESULT 12

ID AAB56582 standard; Protein, 337 AA.

XX AAB56582;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1160.

XX Human, prostate cancer; prostate cancer antigen; detection; diagnosis;
XX neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;
XX vulvovaginal; gastrointestinal; nephropathic; antineoplastic; gynaecological;
XX antibacterial; gene therapy; neural; immune; reproductive; renal;
XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX wound; infectious disease.

XX Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.
XX Rosen CA, Ruben SM;
XX WPI, 2000-567513/55.
DR N-PSDB; AAF15785.

XX Prostate cancer associated gene sequences, referred to as prostate
XX cancer antigens, useful for treatment, prevention, and diagnosis of
XX disorders such as prostate cancer -

XX Claim 11; Page 1566-1567; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX The prostate cancer antigens can have neuroprotective, cytoskeletal,
XX cardioprotective, immunomodulatory, muscular, vulvovaginal, gastrointestinal,
XX nephropathic, antineoplastic, gynaecological and antibacterial activities,
XX and can be used in gene therapy. The prostate cancer antigen
XX polynucleotides may be used for detection of prostate cancer, chromosome
XX identification, as chromosome markers, and for numerous other diagnostic
XX or research purposes. The prostate cancer antigens may be used to treat
XX disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wounds, and infectious diseases. AAF15506 to AAF16514 to
XX AAB57303 represent sequences used in the exemplification of the present
XX invention.

XX Sequence 337 AA;

Alignment Scores:			
Pred. No.:	2.08e-16	Length:	337
Score:	25.00	Matches:	25
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.66%	Indels:	0
DB:	21	Gaps:	0

US-09-896-522-1 (1-1624) x AAB56582 (1-337)

QY 658 ACCTTGCGAAGCCGGCCTTCGAGAGTTCTGCTGCCGACAAAGATATGCCGATGTG 717
Db 263 ThrpheVallyspProAlaPheGluGluPheCysLeuProThrlyblySTyAlaAspVal 282
|||||

QY 718 ATCATCCACGACGAGA 732
Db 283 IleIleProArgGly 287
|||||

RESULT 13

ID AAB41393 standard; Protein, 337 AA.

XX AAB41393;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HOBLP29, SEQ ID NO:2525.

XX Human, ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytoskeletal; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 1p32.

XX Homo sapiens.

PN WO200200677-A1.

KM Musculoskeletal system antigen; cancer; metastasis;
 KM re-vascularisation; thrombosis; arteriosclerosis; mineral content;
 KM cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
 KM post-operative tissue repair; limb regeneration; neuronal growth;
 KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KM AIDS-related complex; chondrocyte growth; bone regeneration;
 KM periodontal regeneration; tissue transport; bone graft; skin aging;
 KM keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
 KM cell growth; organ transplant; cell differentiation; body height;
 KM weight; hair colour; eye colour; skin; percentage of adipose tissue;
 KM pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;
 KM depression; tendency for violence; pain; reproductive capability;
 KM hormone level; endocrine level; appetite; libido; memory; stress;
 KM storage capability; fat content; lipid content; protein content;
 KM carbohydrate content; vitamin content; cofactor content;
 KM nutritional component.
 KM
 KM
 OS Homo sapiens.
 XX
 XX US2002147140-A1.
 PN
 XX
 PD 10-OCT-2002.
 PF
 XX 17-JAN-2001; 2001US-0764877.
 XX
 PR 31-JAN-2000; 2000US-179065P.
 PR 04-FEB-2000; 2000US-180628P.
 PR 28-JUN-2000; 2000US-214886P.
 PR 07-JUL-2000; 2000US-216647P.
 PR 07-JUL-2000; 2000US-216880P.
 PR 11-JUL-2000; 2000US-217487P.
 PR 11-JUL-2000; 2000US-217496P.
 PR 14-JUL-2000; 2000US-218290P.
 PR 26-JUL-2000; 2000US-220963P.
 PR 26-JUL-2000; 2000US-224518P.
 PR 14-AUG-2000; 2000US-224519P.
 PR 14-AUG-2000; 2000US-225267P.
 PR 14-AUG-2000; 2000US-225268P.
 PR 14-AUG-2000; 2000US-225270P.
 PR 14-AUG-2000; 2000US-225447P.
 PR 14-AUG-2000; 2000US-225757P.
 PR 14-AUG-2000; 2000US-225758P.
 PR 22-AUG-2000; 2000US-226868P.
 PR 30-AUG-2000; 2000US-228924P.
 PR 01-SEP-2000; 2000US-229287P.
 PR 01-SEP-2000; 2000US-229343P.
 PR 01-SEP-2000; 2000US-229344P.
 PR 01-SEP-2000; 2000US-229345P.
 PR 05-SEP-2000; 2000US-228509P.
 PR 05-SEP-2000; 2000US-228513P.
 PR 08-SEP-2000; 2000US-231413P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 21-SEP-2000; 2000US-234274P.
 PR 25-SEP-2000; 2000US-234997P.
 PR 27-SEP-2000; 2000US-235834P.
 PR 29-SEP-2000; 2000US-236327P.
 PR 29-SEP-2000; 2000US-236367P.
 PR 29-SEP-2000; 2000US-236368P.
 PR 29-SEP-2000; 2000US-236369P.
 PR 29-SEP-2000; 2000US-236370P.
 PR 02-OCT-2000; 2000US-236802P.
 PR 02-OCT-2000; 2000US-237037P.
 PR 02-OCT-2000; 2000US-237038P.
 PR 02-OCT-2000; 2000US-237039P.
 PR 13-OCT-2000; 2000US-237040P.
 PR 13-OCT-2000; 2000US-239335P.
 PR 20-OCT-2000; 2000US-240960P.
 PR 20-OCT-2000; 2000US-241785P.
 PR 20-OCT-2000; 2000US-241809P.
 PR 01-NOV-2000; 2000US-244617P.
 PR 17-NOV-2000; 2000US-249299P.
 PR 08-DEC-2000; 2000US-251856P.

PR 08-DEC-2000; 2000US-251868P.
 PR 08-DEC-2000; 2000US-251869P.
 XX
 PA (ROSEN/) ROSEN C A.
 PA (RUBEN/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC,
 XX WPI; 2003-128199/12.
 DR N-PSDB; ABX58045.
 XX
 PT Isolated nucleic acid molecules encoding musculoskeletal system
 PT associated polypeptides, useful for detecting disorders, e.g. cancer -
 XX
 PS Claim 11; SEQ ID NO 1422; 321pp; English.
 XX
 CC The invention describes an isolated nucleic acid molecule comprising a
 CC sequence encoding musculoskeletal system associated polypeptides useful
 CC for detecting disorders, e.g., cancer or cancer metastases, in animals
 CC or humans. The nucleic acid: stimulates re-vascularisation of ischaemic
 CC tissues associated with conditions such as thrombosis, arteriosclerosis,
 CC and other cardiovascular conditions; treats wounds due to injuries,
 CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
 CC and limb regeneration; stimulates neuronal growth; can treat and prevent
 CC neuronal damage occurring in certain disorders or neurodegenerative
 CC conditions, such as, Alzheimer's disease, Parkinson's disease, and
 CC AIDS-related complex; stimulates chondrocyte growth, thus they can be
 CC used to enhance bone and periodontal regeneration and aid in tissue
 CC transports or bone grafts; prevents skin aging due to sunburn by
 CC stimulating keratinocyte growth; prevents hair loss, since PGP family
 CC members activate hair-forming cells and promotes melanocyte growth;
 CC stimulates growth and differentiation of hematopoietic cells and bone
 CC marrow cells when used in combination with other cytokines; maintains
 CC organs before transplantation or for supporting cell culture of primary
 CC tissues; induces tissue of mesodermal origin to differentiate in early
 CC embryos; increases or decreases the differentiation or proliferation of
 CC embryonic stem cells, besides, hematopoietic lineage; modulates
 CC mammalian characteristics, such as, body height, weight, hair colour, eye
 CC colour, skin, percentage of adipose tissue, pigmentation, size, and shape
 CC (e.g., cosmetic surgery); modulates mammalian metabolism; changes
 CC mammal's mental state or physical state by influencing biorythms;
 CC cardiac rhythms, depression, tendency for violence, tolerance for pain,
 CC reproductive capabilities, hormonal or endocrine levels, appetite,
 CC libido, memory, or stress; increases or decreases storage capabilities,
 CC fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors
 CC or other nutritional components. This is the amino acid sequence of a
 CC novel human musculoskeletal system antigen.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?DocID=20020147140.
 XX

Sequence 73 AA;

Alignment Scores:

Pred. No.:	3.73e-10	Length:	73
Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.54%	Indels:	0
DB:	24	Gaps:	0

US-09-896-522-1 (1-1624) x ABU12769 (1-73)

Oy 1521 CCACTGAGGAGATTTTCCCTTGTGCTTAAGGACAGAGACGAGCGG 1577
 Db 11 ProlenurGaSpValPheProProCySleuCySleuGlyArgGluAlaArgArg 29

Search completed: November 25, 2003, 07:53:13
 Job time : 78.4988 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 06:52:49 ; Search time 93.8194 Seconds
(without alignments)
8933.708 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 2967
Sequence: 1 gtcggggtcgcctccgcagctc.....ccaggtctgtgctgggagccag 1624

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+np.model -DEV=xlh
-Q/cgn2.1/USPTO.spc01/US09896522/rnat_21112003_184104_2821/app_query.fasta_1.2830
-DB=SPTREMBL_23 -QFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522.qcgn_1.1.182.qrunat_21112003_184104_2821 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMOUT=120 -WARN TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020.5	34.4	201	4 Q96BJ0	Q96BJ0 homo sapien

2	932	31.4	261	11 Q8C476	Q8C476 mus musculus
3	733	24.7	260	5 Q8MRJ1	Q8MRJ1 drosophila
4	572	19.3	486	10 Q9FKS0	Q9FKS0 arabidopsis
5	568	19.1	465	10 Q9LTV6	Q9LTV6 arabidopsis
6	564.5	19.0	419	4 Q8NS24	Q8NS24 homo sapien
7	563	19.0	483	10 Q9LX34	Q9LX34 arabidopsis
8	561.5	18.9	515	5 Q19583	Q19583 caenorhabdi
9	561.5	18.9	555	5 Q9U317	Q9U317 caenorhabdi
10	548	18.5	614	5 Q9V810	Q9V810 drosophila
11	548	18.5	626	5 Q8VQK4	Q8VQK4 drosophila
12	544	18.3	466	10 Q8VYB2	Q8VYB2 arabidopsis
13	529	17.8	469	10 Q8LD95	Q8LD95 arabidopsis
14	508	17.1	1060	10 Q9LFF2	Q9LFF2 arabidopsis
15	507.5	17.1	542	10 Q9LQ32	Q9LQ32 arabidopsis
16	459	15.5	209	16 Q8E0A8	Q8E0A8 streptococ
17	459	15.5	454	3 Q74427	Q74427 schizosacch
18	458	15.4	209	16 Q8ESY5	Q8ESY5 streptococ
19	451	15.2	207	16 Q8CSB2	Q8CSB2 staphylococ
20	441	14.9	209	16 Q8DTG1	Q8DTG1 streptococ
21	439	14.8	419	5 Q9BMX4	Q9BMX4 cryptospori
22	404	13.6	210	16 Q8EPT5	Q8EPT5 oceanodacil
23	403	13.6	213	16 Q8D8F4	Q8D8F4 vibrio vuln
24	390.5	13.2	111	4 Q9BU42	Q9BU42 homo sapien
25	380.5	12.8	111	4 Q92528	Q92528 homo sapien
26	369	12.4	105	11 Q9QYG8	Q9QYG8 rattus norv
27	349	11.8	212	16 Q8EDX4	Q8EDX4 shewanella
28	233	7.9	674	10 Q9C664	Q9C664 arabidopsis
29	212	7.1	643	10 Q9C9B9	Q9C9B9 arabidopsis
30	211.5	7.1	403	10 Q8GRU9	Q8GRU9 oryza sativ
31	210.5	7.1	334	16 Q8YPR9	Q8YPR9 anabaena sp
32	207	7.0	366	10 Q95033	Q95033 odontella s
33	207	7.0	646	10 Q8S3R9	Q8S3R9 oryza sativ
34	205	6.9	448	10 Q8GUE1	Q8GUE1 gaidieria s
35	202	6.8	334	2 Q8LO55	Q8LO55 synchococ
36	202	6.8	334	16 Q8DHN2	Q8DHN2 synchococ
37	199.5	6.7	332	10 P93681	P93681 plectoniv
38	198	6.7	313	16 Q8YU66	Q8YU66 anabaena sp
39	197.5	6.7	333	2 Q9LBU7	Q9LBU7 synchococ
40	197.5	6.7	492	10 Q8RV89	Q8RV89 oryza sativ
41	197	6.6	367	3 Q94642	Q94642 schizosacch
42	195.5	6.6	549	2 Q8GFF2	Q8GFF2 streptomyces
43	189	6.4	572	2 Q68872	Q68872 myxococcus
44	189	6.4	616	4 Q9H6K5	Q9H6K5 homo sapien
45	187.5	6.3	271	10 Q9SPH7	Q9SPH7 beta vulgar

ALIGNMENTS

RESULT 1
Q96BJ0 PRELIMINARY; PRT; 201 AA.
AC Q96BJ0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to uridine-cytidine kinase 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015547; AAH15547.1; -
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PRO0968; URIDINKINASE.
KW KINASE.
SQ SEQUENCE 201 AA; 22790 MW; 0E5F2F00FF7B5363 CRC64;

Alignment Scores:

Pred. No.: 1,33e-72 Length: 201
 Score: 1020.50 Matches: 200
 Percent Similarity: 85.47% Conservative: 0
 Best Local Similarity: 85.47% Mismatches: 1
 Query Match: 34.40% Indels: 33
 Gaps: 1

US-09-896-522-1 (1-1624) x Q96B10 (1-201)

QY 94 ATGGCTTCGGCGGAGGCGGAAGACTCGAGAGCCCGCGCGGAGCGACCGCTCGCAC 153
 Db 1 MetAlaserlaiglygluaspCysgluserProAlaProgluAlaAspArgProHis 20
 QY 154 CAGCGGCTTCCTGATAGGGGTGAGCGCGGCACTCGCCAGCGGAAAGTGCACCTGTGT 213
 Db 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
 QY 214 GAGAAAGTCAATGAGTGTCTGGACAGAACGAGGTGGAACAGCGGAGCGGAAAGTGTGTC 273
 Db 41 GlnuylleuSerGlnleuLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
 QY 274 ATCTGAGCCAGGACAGGTTCTACAGGTCCTGACGCGAGAGCAAGCCAGCCCTTG 333
 Db 61 IleuSerGlnAlaAspArgPheTyrIleValIleuThrAlaGluGlnLysAlaLysAlaLeu 80
 QY 334 AAAGGACAGTCAATTTTGAACATCCAGATCCCTTTGATATATGATTGATGACAGAGACT 393
 Db 81 LysGlyGlnTyrAsnProAspHisProAspAlaPheAspAspLeuMetHisArgThr 100
 QY 394 CTGAAGAACATCGTGGAGGCGAAGACGCTGAGGTCGACCTATGATTTTGTGACACAC 453
 Db 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 120
 QY 454 TCAAGGTACAGAGACCGACGCGTGTCTACCTCGGACGCGTGTCTGTTGAGGCGATC 513
 Db 121 SerArgLeuProGluIleThrValValTyrProAlaAspValValIleuPheGluGlyIle 140
 QY 514 TTGGTGTCTACAGCCAGGAGATCCGGACATGTTCCACCTGCGGCTCTTCTGTGACACC 573
 Db 141 LeuValPheTyrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 574 GACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCGGAGCGTGGCGGAGGAGGACCTG 633
 Db 161 AspSerAspValAlaGluSerArg-Arg----- 169
 QY 634 GAGCAGATTCTGACGAGTACCACTTCGTGAAGCCGCGCTTCGAGAGTTCTGCTG 693
 Db 169 ----- 169
 QY 694 CCGCAAGAGATGTCGGATGTATCATCCACAGAGAGTGCATATNGTTGCCATC 753
 Db 170 --AspLysGluValCysArgCysAspHisProThrArgSerGlyGlnTyrGlyHisG 189
 QY 754 AACCTGATCGTGCAGCAGCATCCAGACATCTTGATG 791
 Db 189 InProAspArgAlaAlaHisProGlyHisSerGluTyr 201

RESULT 2
 Q8C476 PRELIMINARY, PRT, 261 AA.
 AC Q8C476:
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Uridine-cytidine kinase 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;

RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK082837; BAC38646.1; -
 SQ SEQUENCE 261 AA; 29438 MW; 780AA3C298AA8153 CRC64;

Alignment Scores:

Pred. No.: 1.54e-65 Length: 261
 Score: 932.00 Matches: 187
 Percent Similarity: 79.04% Conservative: 28
 Best Local Similarity: 68.75% Mismatches: 41
 Query Match: 31.41% Indels: 16
 Gaps: 4

US-09-896-522-1 (1-1624) x Q8C476 (1-261)

QY 115 GACTGGAGAGAGCCCGCGGAGGCGGACCGTCCGCGAC---CAGCGGCTTCTGATTA 171
 Db 4 AspSerGlnGlnThrLeuGlnAsnHisGlnGlnProAsnGlyGlyGluProPheLeuIle 23
 QY 172 GGGGTGAGCGCGGACCTCGACCGGAAAGTGCACCGTGTGAGAGATCAAGAGTTG 231
 Db 24 GlyValSerGlyGlyThrAlaSerGlyLysSerSerValCysAlaLysIleValGlnLeu 43
 QY 232 CTGGGACAGAACGAGGTGAAACAGCGGACGGAGAGTGTCACTCTGAGCCAGACAG 291
 Db 44 LeuGlyGlnAsnGluValAspTyrHisGlnLysGlnValIleLeuSerGlnAspSer 63
 QY 292 TTCTACAGGTCCTGACGCGGACGAGGCGCAAGGCGCTGAAAGACATGATTTT 351
 Db 64 PheTyrArgValLeuThrSerGlnGlnLysAlaLysAlaLeuLysGluInPheAsp 83
 QY 352 GACCATCCAGATGCTTTGATATGATTTGATGACAGACTCTGAAGACATCGTGAG 411
 Db 84 AspHisProAspAlaPheAspAsnGluLeuIlePheLysThrLeuLysGluIleThrGlu 103
 QY 412 GCGAAACGCGTGGAGGTCGACCTTATGATTTGTACACACTCAAGGTTACAGAGACC 471
 Db 104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGluGluThr 123
 QY 472 ACGGTGTCTACCTCGGAGCGTGTCTGTTGAGGAGATCTGTGTTCTACAGCCAG 531
 Db 124 ValThrIleTyrProAlaAspValValLeuPheGlnGlyIleLeuAlaPheTyrSerGln 143
 QY 532 GAGATCCGGAGACATGTTCCACCTGCGCTTCTGTGAGACACGACTCCGACGTCAGCTG 591
 Db 144 GluValArgAspLeuPheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeu 163
 QY 592 TCTCGAAGATTTCTCGGAGCTG---CGCCGAGGAGGAGACTGAGACAGATTCGAG 648
 Db 164 SerArgArgValLeuArgAspIleSerGlnArgGlyArgAspLeuGlnIleLeuSer 183
 QY 649 CAGTACACACCTTGTGAAGCGGCTCGAGGATTCGTGCTGCGCGCAAGAGAT 708
 Db 184 GlnTyrIleThrPheValIysProAlaPheGlnGluIlePheCysLeuProThrLysValTyr 203
 QY 709 GCCGATGTATCATCCACAGAGAGTGCATATGTTGTCATCACTGATCGTGACG 768
 Db 204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln 223
 QY 769 CACATCAGAGACATTTGTAATGTTGATCTGCAATATGCAACCGAGAGAGTTCATGG 828
 Db 224 HisIleGlnAspIleLeuAsnGly-----Gly 232
 QY 829 CGAGGTACAGCGGACCTTTCTGAGCAGGAGGACCACTGGAGTGTGACCTGTGG 888
 Db 233 LeuSerLysArgGlnThr-----AsnGlyTyrPheAsnGlyIleThrProSerArg 249
 QY 889 AAACGTCACATTTGAGTTCAGACAGACCCAC 924


```

Db      250 LysArgGlnAlaSerGluSerSerArgProHis 261
RESULT 3
ID      Q8MRJ1      PRELIMINARY;      PRT;      260 AA.
AC      Q8MRJ1;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      L013909P.
GN      Cg6364.
OS      Drosophila melanogaster (fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Berkely;
RA      Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA      Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,
RA      George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA      Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA      Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA      Celinker S.;
RL      Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AY119583; AM50237.1; -.
DR      FlyBase; FBgn0039179; CG6364.
DR      InterPro; IPR006083; PRK_URK.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam; PF00485; PRK.1.
DR      PRINTS; PRO0988; URIDINKINASE.
SQ      SEQUENCE 260 AA; 29223 MW; 03CAA6DCA04A16B5 CRC64;

Alignment Scores:
Pred. No.:      1e-49      Length:      260
Score:      733.00      Matches:      142
Percent Similarity:      82.46%      Conservative:      32
Best Local Similarity:      67.30%      Mismatches:      35
Query Match:      24.71%      Indels:      2
DB:      Gaps:      2

US-09-896-522-1 (1-1624) x Q8MRJ1 (1-260)
QY      160 CCCTTCTGATAGGGGTGACCGCGGACCTGCCAGCGGAACTGACCTGTGTGAGAAG 219
DB      27 PrrpPheLeuIleGlyValAlaGlyGlyThrAlaSerGlyLysSerThrValCysLysLys 46
QY      220 ATCATGAGTGTGCGGAGCAAGACGAGTGAACAGCGGAGCGGAAAGGTGATCTCTG 279
DB      47 IleMetGlnIleuGlyGlnAlaGluMetAspHisThrGlnArgGlnValAlaSerIle 66
QY      280 AGCCAGGACAGGTCTTCTCAAGAGTCTGACCGGAGAGCAAGGCAAGGCTTGAAAGCA 339
DB      67 SerGlnAspSerPheThrArgGluLeuThrProAlaGluLysAlaLysAlaGlnLysGly 86
QY      340 CAGTACATTTTGAACCATCCAGATGCTTTGATATGATATTGATGCACAGACTCTGAAG 399
DB      87 LeuPheAsnIleAspHisProAspAlaPheAsnGlnIleuLeuMetLysSerThrLeuGln 106
QY      400 AACATCGTGGAGGCAAAACGAGTGGAGGCGGACCGACCATATTGTCACACACTCA--- 456
DB      107 AsnIleLeuLysGlyHisLysValGluIleProSerLysPheLysArgThrAsnSerLeu 126
QY      457 AGGTTACACAGACCAACGAGTGTCTACCTTCGCGACCTGTGTTGAGGCGATCTTG 516
DB      127 AspPheGluAsnValIleuValIleTyrProAlaAspValValLeuPheGluGlyIleLeu 146
QY      517 GTGTTCTACAGCCAGGAGATCCGGAGACATGTTCCACTCGGCGCTCTTCTGTGACACCGAC 576
DB      147 ValPheTyrPheProLysIleArgGlnLeuPheHisIleMetLysLeuPheValAspThrAsp 166
QY      577 TCCGACGTCAAGGTGTCTCGAAGAGTCTCCGGGAGCTG---CGCCGAGGAGGACCTG 633

```

```

Db      167 ProAspThrArgLeuAlaArgArgValProArgAspIleAsnGlnArgIleArgAspLeu 186
QY      634 GAGCAGATTCTGACGCGATPACACCACTTCTGTGAACCGCGCTTGAGGAGTTCCTTG 693
DB      187 AspAlaValLeuThrGlnTyrMetThrPheValLysProAlaPheGluIlePheCysSer 206
QY      694 CCGACAAAGAAATGATGCCATGTGATCATCCACGAGAGTGAGCAATATGTTGCCATC 753
DB      207 ProThrLysLysPheAlaAspValIleIleProArgGlyAlaAspAsnThrValAlaIle 226
QY      754 AACCTGATGTCAGACATCCAGACATCCAGACATCTTG 786
DB      227 AspLeuIleValAlaHisIleIleGlyIleLeu 237

RESULT 4
ID      Q9FKS0      PRELIMINARY;      PRT;      486 AA.
AC      Q9FKS0;
DT      01-MAR-2001 (TEMBLrel. 16, Created)
DT      01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      Uridine kinase-like protein.
GN      AT5G40870.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Columbia;
RX      MEDLINE=98344145; PubMed=9679202;
RA      Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA      Tabata S.;
RT      "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT      features of the regions of 1,381,565 bp covered by twenty one
RT      physically assigned pl and TAC clones.";
RL      DNA Res. 5:131-145 (1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA      Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,
RA      Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,
RA      Palm C.J., Shinn P., Southwick A., Tripp M.G., Wu T., Davis R.W.,
RA      Ecker J.R., Theologis A.;
RT      "Arabidopsis Open Reading Frame (ORF) Clones.";
RL      Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AB011477; BAB11349.1; -.
DR      EMBL: BT002336; AM86169.1; -.
DR      HSSP; Q26998; 1BD3.
DR      InterPro; IPR006083; PRK_URK.
DR      InterPro; IPR000836; PRTtransferase.
DR      InterPro; IPR00764; Uridine_kin.
DR      Pfam; PF00156; Priboylcran.1.
DR      Pfam; PF00485; PRK.1.
DR      PRINTS; PRO0988; URIDINKINASE.
DR      TIGRPFAM; TIGR00235; udk.1.
KW      kinase; transferase.
SQ      SEQUENCE 486 AA; 54430 MW; 50DA1CE89346FB54 CRC64;

Alignment Scores:
Pred. No.:      7.42e-37      Length:      486
Score:      572.00      Matches:      123
Percent Similarity:      61.38%      Conservative:      55
Best Local Similarity:      42.41%      Mismatches:      76
Query Match:      19.28%      Indels:      36
DB:      Gaps:      8

US-09-896-522-1 (1-1624) x Q9FKS0 (1-486)
QY      7 TCGCTCCGACCTCGAGCGGTGGCGGCGCGCGGCGCGGAGGAGGCGGCGGCGG 66

```

```

Db      32 SerProProAenSerSerVal-----38
Qy      67 ACCGGATGCGCGGAGCGGAGCCGAGATGCTTGGCGGAGCGGAGACTGCGGAGAC 126
Db      39 -----ValSerSerLeuArgSerAlaValSerSerSerProSerSerSerAapPro 56
Qy      127 CCGCGCGCGGAGCGGAGCCGCGGAGCGGAGCTTCCGATGAGGGGTGAGGGGGGCGG 186
Db      57 GluAlaPro-----LysGlnProPheIleIleGlyValSerGlyGly 70
Qy      187 ACTGCGAGCGGAGATGCGACCGTGTGTGAGAGATCATGAGATTGCTGGACAGAGAG 246
Db      71 ThrAlaSerGlyLysThrThrValCysAspMetIleIleGlnGlnLeu-----86
Qy      247 GTGAAACAGCGGAGCGGAGAGTGTCTCTGAGACCGAGAGTTTACAGAGTCTG 306
Db      87 -----HisAspHisArgValValLeuValAsnGlnAspSerPheTyArgGlyLeu 103
Qy      307 ACGGAGAGCGGAGGCGGAGCGCTTGAAGAAGACATGATTTTGAACCATCGAGAGCC 366
Db      104 ThierGlnGlnLeuGlnArgVal-----GlnGlnTyAsnProAspHisProAspHis 121
Qy      367 TTTGATATGATTTGATGACAGACACTCTGAAGAATCTGAGGCGGAGGAGGAG 426
Db      122 PheAspThrGlnGlnLeuLeuHisGlyAlaGlnThrLeuLysSerGlyGlnProTyGln 141
Qy      427 GTGCGGAGCTTGTATTTGTGACACATGAGTTTACAGAGAC--ACGGTGTCTAC 483
Db      142 ValProIleTyArgPheLysThrHisGlnArgArgSerAspThrPheArgGlnValAsn 161
Qy      484 CCGCGGAGCGGAGGCTTGTGAGGAGCTTGTGAGGCTTCTGAGCGGAGAGAGAG 543
Db      162 AlaSerAspValIleIleLeuGlnGlyLeuValPheHisAspSerArgValAlaArgAsn 181
Qy      544 ATGTTCACCTGCGCGCTTCTGTGACACCGCATCCGAGCTGTCTGGAAGATT 603
Db      182 LeuMetAspMetLysIlePheValAspThrAspHisArgValAlaArgLeuAlaArgGlyLe 201
Qy      604 CTCGGGAGC--GTGCGCGGAGGAGGAGCGCTGAGAGCATTTCTGAGCGGAGTACCAACC 660
Db      202 ArgArgAspThrValGlnArgGlyArgAspValAsnSerValLeuGlnGlnTyArgLys 221
Qy      661 TTGCGTGAAGCGGCGCTTGTGAGAGATGCTGCGCGGAGCAAGAAGATGCGGAGTATC 720
Db      222 PheValLysProAlaPheAspAspPheValLeuProSerLysTyArgAlaAspValIle 241
Qy      721 ATCCGACGAGAGTGAACAATATGTTGCCATCAACTGATCTGTGACAGACATCCAGAGC 780
Db      242 IleProArgGlyGlyAspAsnHisValAlaValAspLeuIleThrGlnHisIleHisThr 261
Qy      781 ATTCTG--AATGCGACATCTGCAAAATGAGCAGGAGAGGTCCATGGCGGAGCTAC 837
Db      262 LysLeuGlnGlnHisAspLeuCysLysIleTy-----ProAsnValTyArgValIle 278
Qy      838 AAGCGAGCTTTCTGAGCGGAGGAGCAGC 867
Db      279 GlnSerThrPheGlnIleArgGlyMetHis 288

```

RESULT 5

```

Q9LTY6 PRELIMINARY; PRT; 465 AA.
AC Q9LTY6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Uridine kinase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_taxid=3702;
RN
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819229;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL; AB024028; BAA95720.1; -
DR HSSP; Q26998; 1BD3
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; URK1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KW kinase.
SQ SEQUENCE 465 AA; 52125 MM; 4CD82DB9059E7168 CRC64;

```

Alignment Scores:

	Pred. No.:	Length:	Matches:
Score:	1.52e-36	465	568.00
Percent Similarity:	66.94%		Conservative: 51
Best Local Similarity:	46.37%		Mismatches: 68
Query Match:	19.14%		Indels: 14
DB:	10		Gaps: 5

```

US-09-896-522-1 (1-1624) x Q9LTY6 (1-465)
Qy      130 GCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 189
Db      19 SerProSerLeuProAlaProLeuLysGlnProPheValIleGlyAlaGlyGlyThr 38
Qy      190 GCCAGCGGAGAGCGGAGCGGAGTGTGAGAGATCATGAGTGTGCGGAGAGAGAGT 249
Db      39 AlaSerGlyLysThrThrValCysAspMetIleMetSerGlnLeu-----53
Qy      250 GAACAGCGGAGCGGAGAGTGTGATCTGAGCGGAGAGAGTGTCTACAGAGTCTGAGC 309
Db      54 -----HisAspLysArgValValLeuValAsnGlnAspSerPheThrHisSerLeuThr 71
Qy      310 GCAGACGAGAGGCGCAAGCGCTTGAAGAAGACATCAATTTGACATCCAGATGCTT 369
Db      72 LysGlnLysLeuAsnLysVal-----HisGlnTyAsnAspHisProAspAlaPhe 89
Qy      370 GATATGATTTGATGACAGACTGTAAGAATCATCTGAGGAGGAGCAAAACGGTGAAGT 429
Db      90 AsnThrGlnValLeuLeuSerCysMetGlnLysLeuArgSerGlyGlnProValAsnIle 109
Qy      430 CCGAGCTATGATTTGTGACACACTCAAGTTACCAAGACCAAGCGGTGCTTACCTCGG 489
Db      110 ProSerTyAspPheLysIleHisGlnSerIleGlnLysSerSerProValAsnProGly 129
Qy      490 GAGCTGTTCTGTTTGAAGGAGCATCTGAGTGTCTTACAGCGAGAGATCCGGGAGATT 549
Db      130 AspValIleLeuGlnGlyLysLeuValLeuAsnAspProArgValAlaArgAspLeuMet 149
Qy      550 CACCTGCGCTCTTGTGACACCGACTCGAGCTGAGCGTGTCTGAGAGAGTCTCCGG 609
Db      150 AsnMetLysIlePheValAspThrAspHisArgValAlaArgLeuSerArgGlyGlnArg 169
Qy      610 GAC--GTGCGCGGAGGAGGAGCGGAGAGATTTGACGAGTACCAACCTTCTGTG 666
Db      170 AspThrValGlnArgGlyArgAsnIleGlnAsnValLeuGlnGlnTyThrLysPheVal 189
Qy      667 AACCGCGCTTGAAGAGTTCCTGCGGAGCAAAAGATGAGTGTGATGATCATCCCA 726
Db      190 LysProSerPheAspLysGlnIleGlnProSerMetLysTyArgAlaSerIleIlePro 209
Qy      727 CGAGAGTGAACAATATGTTGCTCATCAACTGATCTGTGAGAGACATCCAG--GACATT 783

```

Db	210	AGGGLGylGAspAsnAspValAlaIleAdePheIleIleValGlnHisIleLeaGlyThrIysLeu	229
QY	784	CTGAATGGTGCATCTGCAATATGGCACCGAGAGGGTCCAAATGGCGCGAGACTTCAAGCGG	843
Db	230	CysGlnHisAsnLeuCysIysIleTyr-----SerAsnIlePheIleIleSerSer	246
QY	844	ACCTTTTCGACCGCAGGAGCCAC	867
Db	247	ThrPheGlnIleIysGlyMetHis	254
RESULT 6			
ID	Q8N524	PRELIMINARY;	PRT; 419 AA.
AC	Q8N524;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Similar to uridine kinase-like 1.		
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RA	Straussberg R.;		
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC033078; AAH33078.1; -		
DR	InterPro; IPR006082; PRK.		
DR	InterPro; IPR006083; PRK_URK.		
DR	InterPro; IPR007074; Uridine_kin.		
DR	Pfam; PF00485; PRK; 1.		
DR	PRINTS; PRO0478; PHRI-BKINASE.		
DR	PRINTS; PRO0988; URIDINKINASE.		
DR	TIGRFAMs; TIGR00235; udk; 1.		
KW	Kinase.		
SQ	SEQUENCE 419 AA; 45972 MW; 1D1F1E75305AD25 CRC64;		
Alignment Scores:			
	Pred. No.:	2,81e-36	Length: 419
	Score:	564.50	Matches: 148
	Percent Similarity:	46.70%	Conservative: 57
	Best local Similarity:	33.71%	Mismatches: 125
	Query Match:	19.03%	Indels: 109
	DB:	4	Gaps: 13
US-09-896-522-1 (1-1624) x Q8N524 (1-419)			
QY	7	TCGCCTCGACCTCGGCGCTGGGCGGCGCGCGCGCGCGGAGGAGGCGCGCGCGG	66
Db	16	SerProProThrAla-----ArgAspThrProGlyArg-----	26
QY	67	ACCCGATGCGCGGAGCGGAGCGGAGCGCGATGCGCTTCGCGCGGAGCGGAGAC-----	117
Db	27	-----GlnAlaGlnIysSerGluThrAlaCysGlnAspArgSerAsn	40
QY	117	-----	117
Db	41	AlaGlnSerLeuAspArgLeuProProValGlyThrGlyArgSerProArgIysArg	60
QY	118	-----TGGCAGAGCCCCGCGCGG-----	135
Db	61	ThrThrSerGlnCysIysSerGluProProLeuLeuArgThrSerIysArgThrIleTyr	80
QY	136	GAGGCGGACCCGCG-----CACACGCGGCGCCTTCCTG	168
Db	81	ThrAlaGlyArgProProTyrAsnGlnHisGlyThrGlnSerIysGlnAlaPheAla	100
QY	169	ATAGGCGTGAAGCGCGGCGCATGCGACGCGGAGATCGACCGTGTGTGAAGAATCATGAG	228
Db	101	IleGlyLeuGlyGlyGlySerIAsaSerGlyIysThrThrValAlaArgMetIleIleGlu	120

OY	229	TTGGTGGGACAGAAAGAGTGGAAACAGCGCAGCGGAAAGCGTACCTTGAGCGCAGAC	288
Db	121	AlaLeu-----AspValProTrp-----ValValLeuLeuSerMetasp	133
OY	289	AGGTTTACAGAGTCTCTGACGGCAGAGCAGAGGCAAGGCGCTTGAAAGCAGTACAAAT	348
Db	134	SerHeuTrpValLeuThrGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	153
OY	349	TTTGAACATCCAGATGCCCTTTGATTAATGATTTGATTCACAGAGACTCTGAAGACATCGTG	408
Db	154	PheAspHisProAspAlaPheAspPheAspLeuLeuLeuSerThrLeuValLeuValLeuVal	173
OY	409	GAGGCGAAAGCGGAGAGTCCGACACCTATGATTTTGACACACCTCAAGGTTACCAAG	468
Db	174	GlnGlyLysSerValLysValProIleTrpAspPheThrThrHisSerArgLysLeuAsp	193
OY	469	ACCAACGGGAGTCACTCCCTGCGACAGTGGTGTCTTTGAGGCGCACTTGTTCTTACAGC	528
Db	194	TrpLysThrLeuTrpGlyAlaAsnValIleLeuPheGluGlyIleMetAlaPheAlaAsp	213
OY	529	CAGAGATCCGGACATGTTCCACTCGCGCTTTGTCGACACAGCTCCGACGTCCAGG	588
Db	214	LysThrLeuLeuGluLeuLeuAspMetLysIlePheValAspThrAspSerAspIleArg	233
OY	589	CTGTCTCGAAGAGTTCCTCCGGAGCTG---CGCCGAGGAGGAGACTGAGACATTCGTG	645
Db	234	LeuValAlaArgAlaLeuArgAlaAspIleSerGluLysGlyAlaArgAspIleGluGlyValIle	253
OY	646	ACGCAAGTACACCACTTCGTGACACCGGCGCTTGACGAGTCTGCTCGCAGCAAAAG	705
Db	254	LysGlnTrpLeuLysPheValLysProSerPheAspGlnTrpIleGlnProThrMetArg	273
OY	706	TATCCGATGTGATCATCCACGAGAGTGGACAAATATGTTGCGATCAACCTGATCGTG	765
Db	274	LeuAlaAspIleValAlaProArgGlySerGlyLysThrValAlaIleAspLeuIleVal	293
OY	766	CAGCAATCCAGGACATTTGATATGGTACATCTGCAATATGGACACCGAGAGGCTCAAT	825
Db	294	GlnHisValHisSerGlnLeu-----GluGluGlyCysAla	305
OY	826	GGGGGAGGCTCAAGCGGACCTTTCTTGAGCCA---GGGACACCACTGGGATGCTGACC	882
Db	306	GlyLeuGlyThrProValProGlnAlaAlaProAspAlaGluLysProGluGlnHisAla	325
OY	883	TCTGCGAAACGGTCACTTGGAGTCCAGCAGCAGCAGCCCACTGAGGGGCTGCCAGGCT	942
Db	326	AlaGlyThrGlyHisValHisHisHisGlnValAlaArgAlaHisLeuGlyThrGlyArgPro	345
OY	943	CAGGCGAGGTCCTCCCGCGGCGATGTGTGTTCAAGGACTGAGCTGGGAGCGCCACCCA	1002
Db	346	ArgAlaAlaArgAlaProAlaGln-----	352
OY	1003	CACCCACTGCTTCTCTCGGGGCAACCCAGGGAGTGTTCAGCAGAGCC-----	1055
Db	353	-----LeuLeuAlaProGlnGlyGlnGlyArgGlnSerArgValHis	367
OY	1054	TTCTCTACTCAGAGTGGAAACTCAGATGTGTCACTCAGACTCAACTGTCTGGAGACTG	1113
Db	368	LeuLeuLeuGlnGlnThrAspAlaAlaAlaHisArgAlaArgAlaLeuLeuProAlaLeu	387
OY	1114	ACAGGCGTTCCTGAGGTTTTCAGGCCACTTAGGCTGTTCGGTTTAAAGATCCCTCT	1170
Db	388	SerGly-----LeuArgGlyThrAspProAla	396
RESULT 7			
O9LK34			
AC	O9LK34	PRELIMINARY; PRT; 483 AA.	
DT	01-OCT-2000 (TREMBlrel. 15, Created)		
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)		
DT	01-MAR-2003 (TREMBlrel. 23, Last annotation update)		
DS	Uridine kinase-like protein.		
	Arabidopsis thaliana (Mouse-ear cress).		

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RA MEDLINE=20363099; PubMed=10907853;
 RX Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety PL,
 RT TAC and BAC clones";
 RL DNA Ref. 7:217-221(2000).
 DR EMBL; AP000381; BAB02114.1; -.
 DR HSSP; Q26998; 1BD3.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 KW Kinase.
 SQ SEQUENCE 483 AA; 54210 MW; 2862F7AED187020F CRC64;
 Alignment Scores:
 Pred. No.: 3.85e-36 Length: 483
 Score: 563.00 Matches: 117
 Percent Similarity: 67.08% Conservative: 46
 Best Local Similarity: 48.15% Mismatches: 64
 Query Match: 18.98% Indels: 16
 DB: 10 Gaps: 7
 US-09-896-522-1 (1-1624) x Q9LK34 (1-483)
 QY 148 CCGGACCCAGCCGCTTCTGATAGGCTGAGCGGCGGACCTGCAGCGGGAAGTGCAC 207
 Db 59 Protheglin---ProtheglinIleGlyValIhrGlyGlyThrAlaSerGlyValThrThr 77
 QY 208 GTGTGTGAAAGATCAGTGGCTGCGGACGAAACGAGTGAACGCGGACCGGAG 267
 Db 78 ValCysAspMetIleIleGlnGlnLeu-----HisAspHisArg 90
 QY 268 GTGTGATCTCTGAGCGGACGAGTTCACAGGCTCTGAGCGGACGAAAGCCACG 327
 Db 91 IleValIleuValIleuGlnAspSerPheTyrArgGlyLeuThrSerGlnIleuGlnHis 110
 QY 328 GCCTTGAAGGACGATCAATTTGACCATCCAGATCGCTTGAATGATTGATGCAC 387
 Db 111 Val-----GlnGluTyrAsnPheAspHisProAspAlaPheAspThrIleGlnIleu 128
 QY 388 AGAAGCTGAAAGAACATGCTGAGGCAAAACGCTGAGGCTGCGGACCTGATTTGTG 447
 Db 129 HisCysValIleuIleuLeuSerGlyGlnProTyrGlnIleProIleTyrAspPheLys 148
 QY 448 ACAAGCTCAAGGTTACCGAGACCC---ACGGTGTCTACCTCGGACGCTGCTCTTT 504
 Db 149 ThrHisGlnArgLysValAspAlaPheArgGlnValAsnAlaCysAspValIleIleu 168
 QY 505 GAGGACATCTTGTGTTCTTACAGCCAGAGATCCGGGACATGTTCCACCTGCGCTTTC 564
 Db 169 GluGlyIleLeuValPheHisAspSerArgValArgAspLeuMetAsnMetLysIlePhe 188
 QY 565 GTGACACCGACCTCGACGCTGCTGTCTGAAAGAGTTCTCCGGAC--GTGCGCGGA 621
 Db 189 ValAspThrAspAlaAspValArgLeuAlaArgArgIleArgArgAspThrValGluArg 208
 QY 622 GGGGAGGAGCTGAGGACATTCGACGAGTACACCACTTCGGAAGCGGCTTGGAG 681
 Db 209 G1yArgAspValAspSerValIleuGlnIleTyrAlaLysPheValLysProAlaPheAsp 228

QY 682 GAGTTCGCTCGCGCAAAAGATATGCCGATGTATCATCCACGAGAGTGCACAT 741
 Db 229 AspPheValIleuProSerLysTyrAlaAspValIleIleProArgLysGlyAsn 248
 QY 742 ATGTTGCCATCAACTGATCGTCAGACATCCAGACATTCTG--AATGTGACATC 798
 Db 249 HisValAlaValAspLeuIleValGlnHisIleThrLysLeuGlyGlnHisAspLeu 268
 QY 799 TGCAGATGCCACGAGAGGATCCAAATGGCGGAGCTACAGCGGACCTTTCTGAGCCA 858
 Db 269 CysLysIleTyr-----ProAsnValPheValIleGluThrPheGlnIleArg 285
 QY 859 GGGGAGCCAC 867
 Db 286 GlyMetHis 288
 RESULT 8
 Q19583 ID Q19583 PRELIMINARY; PRT; 515 AA.
 AC Q19583;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE F19B6.1a protein.
 GN F19B6.1 OR F19B6.1A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thomas K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z69635; CA93459.1; -.
 DR HSSP; Q26998; 1BD3.
 DR WormPep; F19B6.1a; CE05666.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK URK.
 DR InterPro; IPR000764; Uridine_kin.
 DR Pfam; PF00485; PRK; 1.
 DR PRINTS; PRO0478; PHRIBLKINASE.
 DR PRINTS; PRO0988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk; 1.
 SQ SEQUENCE 515 AA; 58452 MW; 9584D947A2D3B5C5 CRC64;
 Alignment Scores:
 Pred. No.: 5.15e-36 Length: 515
 Score: 561.50 Matches: 125
 Percent Similarity: 63.10% Conservative: 46
 Best Local Similarity: 46.13% Mismatches: 62
 Query Match: 18.92% Indels: 38
 DB: 5 Gaps: 6
 US-09-896-522-1 (1-1624) x Q19583 (1-515)
 QY 12 TCCGACCTCGGCGCTGG---GGGCGCGCGCGGCGCGGGAAGCGGCGGCGGAC 68
 Db 14 SerProArgAlaIleGlyCysArgThrArgArgArgThrMetSerGlyValArgAlaGlu 33
 QY 69 CCGATGCGCGGAGCGGAGCGCGAGATGCTTCGCGGAGGCGAAGACTGCGGAGGCC 128
 Db 34 -----GCGCGCGCGGAGCGCGAGCGCGAGCGCGAGCGCGAGCGCGAGCGCG 153
 QY 129 GCGCGCGGA-----GGCGCGCGCTCGGAC-----HisIleuLeuThrThr 39


```

DE LD03595P.
GN CG4798.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Prounenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129436; AM76178.1; -.
DR FlyBase; FBgn0034213; CG4798.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAms; TIGR00235; udk; 1.
SQ SEQUENCE 626 AA; 70001 MW; 9DBEC0A7331A6E93 CRC64;

Alignment Scores:
Pred. No.: 6,43e-35 Length: 626
Score: 548.00 Matches: 116
Percent Similarity: 60.37% Conservative: 47
Best Local Similarity: 42.96% Mismatches: 61
Query Match: 18.47% Indels: 46
Gaps: 6

US-09-896-522-1 (1-1624) x Q8MQK4 (1-626)
OY 16 ACCTCGCGCGTGGCGGCGC-----GCCGCGCGCGCGGAA----- 51
Db 144 ThrtHrSerIleIysngIntrHrThrAlAsnProSerIuYsIleIleAglA 163
OY 52 -----GGGCGCGCGCGCGGGAACCCGATCGCGGAGCGCGAG 87
Db 164 AsnaAsnArgThrIleTyrThrAlaGlyArgProProTyrIAsnGlyAlaGlnGln 183
OY 88 GCCGAATGCGTTCGGCGGAGCGGAAGACTCGAAGACCCCGCGCGAGCCGACCGT 147
Db 184 ValGlu----- 185
OY 148 CGGCACACAGCGCCCTTCCTAGAGGGGTGAGCGCGGACATGCGAGCGGGAAGTGAC 207
Db 186 -----ProPheValIleGlyIleCyseGlyIleSerAlaSerGlyIlyThr 201
OY 208 GTGTGTGAGAAGATCATGAGTTGTGGGACAGAACGAGGTGAGAACGCGGACGAG 267
Db 202 ValAlaGluIysIleIleGluSerIeu-----AspAlaProTyr----- 214
OY 268 GTGTGATCTGTGACGACGAGCATGTTCTTACAGAGTCTGACGCGACAGCAAGCC 327
Db 215 ValThleuIeuSerMetAspCyseIlyIysIleIeuAsnGlnIuysGlnIsgluGln 234
OY 328 GCCTTGAAGAGACATGATTTGACCATCCAGATGCGCTTATATGATTTGATGCAC 387
Db 235 AlaIeuIleAsnGlnIuysIAsnProAspAlaPheAspIleGluLeuLeu 254
OY 388 AGGACTCTGAAGAACAATCGTGAAGGCAAAACGTTGAGGTGCGGACCTTATGATTT 447
Db 255 AspValIeuThrIysIeuIysGluGlyArgIysValGluValProValIyAsnAsp 274
OY 448 ACACTCAAGCTTACAGACGAGCGGTGCTTACCTCGGACGAGTTCGTTGTTGAG 507
Db 275 ThHieGlyArgIsgluSerGlnThrIlysthMetIyGlyAlaAsnValIleIlePheGlu 294
OY 508 GGCATCTGTGTTTCTACAGCGAGATCGGGAATGTTTCAACCTGCGCGCTTCGTG 567

```

```

Db 295 GlyIleIeuThrPheIisSerProGluValIeuIysIleIeuAspMetIysIlePheVal 314
OY 568 GACACCGACTCCGACGTACAGGCTGTCTCGAAGATTTCTCGGACGCTG---CGCCAGAG 624
Db 315 AspThrAspProAspIleArgIleuAlaArgAluArgAspIleSerGlnArgGly 334
OY 625 AGGAGCTCGACAGATTCGTACGACGTACACACCTTGTAAGCGGCGCTTCGAGAG 684
Db 335 ArgAspIeuIysGlyValIeuIysGlnIyIleuAsnMetValIysProSerTyrCyAsn 354
OY 685 TTCTGCTCCCGACAAAGAGTATGCCGATGTATCATCCACGAGAGTGAATATATG 744
Db 355 TyrIleAlaProThrMetAlaIshAlaAspIleIleValProArgGlyIyIAspAsnIys 374
OY 745 GTTCCCATCAACCTGATCGTGACGACATC 774
Db 375 ValAlaIleIshIeuIleValGlnIshVal 384

RESULT 12
ID Q8YB2 PRELIMINARY; PRT; 466 AA.
AC Q8YB2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative uracil phosphoribosyltransferase.
GN A7IG55810.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxId=702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banb J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Tortum M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinzaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AtIG55810 (GI:15222778).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Heuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Tortum M., Wu H.C., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A7072218; AAL60039.1; -.
DR EMBL; AY129446; AAM67479.1; -.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAms; TIGR00235; udk; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 466 AA; 52443 MW; 3FA6783CB49727F7 CRC64;

Alignment Scores:
Pred. No.: 1.23e-34 Length: 466
Score: 544.00 Matches: 116
Percent Similarity: 63.16% Conservative: 52
Best Local Similarity: 43.61% Mismatches: 78
Query Match: 18.34% Indels: 20
Gaps: 8

```


